

Aus der Medizinischen Klinik und Poliklinik IV
Klinikum der Ludwig-Maximilians-Universität München

Direktor: Prof. Dr. med. Martin Reincke

**Analyse und Validierung eines Messinstruments kognitiver
Prüfungsangst im Kontext des ersten medizinischen
Staatsexamens**

Dissertation

zum Erwerb des Doktorgrades der Medizin

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München

vorgelegt von

Alexandra Gabriele Stefan

aus

Fürstenfeldbruck

2019

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter:	PD Dr. med. Matthias Angstwurm, MME
Mitberichterstatter:	Prof. Dr. med. Martin Fischer Prof. Dr. med. Jörg Schelling
Mitbetreuung durch die promovierte Mitarbeiterin:	Dr. med. Dr. jur. Christina M. Berchtold, MME
Dekan:	Prof. Dr. med. dent. Reinhard Hickel
Tag der mündlichen Prüfung:	17.12.2020

Meinen Eltern gewidmet.

Inhaltsverzeichnis

1 EINLEITUNG	5 -
1.1 PRÜFUNGSANGST	5 -
1.1.1 Definition und Konstrukt, die „Zwei-Komponenten-Theorie“	5 -
1.1.2 Entstehung.....	6 -
1.1.3 Stabilität	6 -
1.1.4 Negative Auswirkungen	7 -
1.1.4.1 Prüfungsangst und Prüfungsleistung	7 -
1.1.5 Reduktion.....	8 -
1.1.5.1 Prüfungssimulationen	9 -
1.1.6 Geschlechtsspezifische Unterschiede.....	9 -
1.1.7 Messinstrumente von Prüfungsangst	10 -
1.1.7.1 „Test Anxiety Inventory“ (TAI) von Spielberger	10 -
1.1.7.2 „Reactions to Tests“ (RTT) von Sarason	10 -
1.1.7.3 „Cognitive Test Anxiety Scale“ (CTAS) von Cassady & Johnson.....	11 -
1.2. MEDIZINSTUDIERENDE & MEDIZINSTUDIUM	11 -
1.2.1 Der Erste Abschnitt der Ärztlichen Prüfung (Physikum).....	11 -
1.2.2 Medizinstudierende als Risikogruppe	12 -
2 FRAGESTELLUNG UND ZIELSETZUNG.....	13 -
3 MATERIAL & METHODEN	14 -
3.1 AUFBAU DER STUDIE	14 -
3.1.1 Studiendesign	14 -
3.1.1.1 Durchführung von Examenssimulationen	15 -
3.1.2 Datenschutz und Datennutzung	15 -
3.1.3 Ethikvotum	16 -
3.2 MESSINSTRUMENTE	16 -
3.2.1 „Cognitive Test Anxiety Scale“ (CTAS)	16 -
3.2.1.1 Übersetzung ins Deutsche.....	16 -
3.2.2 Zusatzmaterial	17 -
3.3 STICHPROBEN.....	20 -
3.4 STATISTISCHE AUSWERTUNG	22 -
3.4.1 Instrumente und Programme.....	22 -
3.4.2 Testverfahren.....	22 -
3.4.2.1 Tests zur Überprüfung von Unterschiedshypothesen	22 -
3.4.2.2 Korrelationen	22 -
3.4.2.3 Signifikanz.....	23 -
3.4.3 Item- und Testanalyse, Validierung von G-CTAS	23 -

3.4.3.1 Deskriptive Statistik und Itemanalyse	- 23 -
3.4.3.2 Exploratorische Faktorenanalyse (EFA)	- 24 -
3.4.3.3 Skalenreliabilität	- 24 -
3.4.3.4 Inhaltsvalidität.....	- 24 -
3.4.4 Verteilungsanalyse kognitiver Prüfungsangst.....	- 24 -
3.4.4.1 Geschlechtsunterschiede	- 24 -
3.4.5 Selektionsbias.....	- 25 -
3.4.6 Stabilität kognitiver Prüfungsangst	- 25 -
3.4.7 Ausmaß von Sorgen vor dem Physikum und mögliche Zusammenhänge	- 25 -
3.4.8 Examenssimulation	- 25 -
3.4.8.1 Zielgruppe der Examenssimulation.....	- 25 -
3.4.8.2 Effekte der Examenssimulation	- 26 -
4 ERGEBNISSE	- 27 -
4.1 ITEM- UND TESTANALYSE, VALIDIERUNG VON G-CTAS	- 27 -
4.1.1 Deskriptive Statistik und Itemanalyse	- 27 -
4.1.2 Exploratorische Faktorenanalyse (EFA)	- 32 -
4.1.3 Skalenreliabilität.....	- 34 -
4.1.4 Inhaltsvalidität	- 34 -
4.2 VERTEILUNGSANALYSE KOGNITIVER PRÜFUNGSANGST	- 36 -
4.2.1 Geschlechtsunterschiede.....	- 37 -
4.3 SELEKTIONSBIAS	- 38 -
4.4 STABILITÄT KOGNITIVER PRÜFUNGSANGST	- 40 -
4.5 AUSMAß VON SORGEN VOR DEM PHYSIKUM UND MÖGLICHE ZUSAMMENHÄNGE ...	- 41 -
4.6 EXAMENSSIMULATIONEN	- 45 -
4.6.1 Zielgruppe der Examenssimulation	- 45 -
4.6.2 Effekte der Examenssimulation	- 47 -
5 DISKUSSION.....	- 53 -
5.1 DISKUSSION DES MATERIALS UND DER METHODEN.....	- 53 -
5.1.1 Zusammenfassung des Materials und der Methoden	- 53 -
5.1.2 Aufbau der Studie	- 54 -
5.1.2.1 Studiendesign.....	- 54 -
5.1.2.1.1 Durchführung als online-Studie	- 54 -
5.1.2.1.2 Durchführungs- und Befragungszeitpunkt	- 54 -
5.1.2.1.3 Sponsoring der Examenssimulationen durch einen Verlag	- 55 -
5.1.2.2 Datenschutz	- 55 -
5.1.3 Messinstrumente.....	- 56 -
5.1.3.1 Verwendung von CTAS	- 56 -
5.1.3.2 Übersetzung ins Deutsche.....	- 57 -

5.1.3.3 Zusatzmaterial.....	- 57 -
5.1.3.4 Antwortformate	- 58 -
5.1.4 Stichproben.....	- 58 -
5.1.4.1 Rücklaufquoten und Stichprobenumfänge	- 58 -
5.1.4.2 Geschlechterverteilung.....	- 60 -
5.1.4.3 Stichprobenzuteilung.....	- 60 -
5.1.5 Statistische Auswertung.....	- 61 -
5.1.5.1 Item- und Testanalyse, Validierung von G-CTAS	- 61 -
5.1.5.1.1 Deskriptive Statistik und Itemanalyse.....	- 61 -
5.1.5.1.2 EFA.....	- 61 -
5.2 DISKUSSION DER ERGEBNISSE	- 63 -
5.2.1 Zusammenfassung der Ergebnisse.....	- 63 -
5.2.2 Item- und Testanalyse, Validierung von G-CTAS.....	- 64 -
5.2.2.1 Deskriptive Statistik und Itemanalyse, EFA	- 64 -
5.2.2.2 Überprüfung der Validität von G-CTAS.....	- 66 -
5.2.2.3 Vorteile und Einschränkungen bezüglich der Nutzung von G-CTAS.....	- 66 -
5.2.3 Verteilungsanalyse kognitiver Prüfungsangst, Analyse von Geschlechtsunterschieden.....	- 67 -
5.2.4 Selektionsbias.....	- 68 -
5.2.5 Stabilität kognitiver Prüfungsangst	- 69 -
5.2.6 Ausmaß von Sorgen vor dem Physikum und mögliche Zusammenhänge.....	- 69 -
5.2.7 Examenssimulationen	- 70 -
5.2.7.1 Zielgruppe der Examenssimulation.....	- 70 -
5.2.7.2 Effekte der Examenssimulation	- 71 -
5.2.7.2.1 Kognitive Prüfungsangst	- 71 -
5.2.7.2.2 Sorgen vor schriftlichem und mündlichem Physikum.....	- 71 -
5.2.7.2.3 Lernaufwand	- 72 -
5.2.7.2.4 Prüfungsleistungen	- 72 -
6 FAZIT & AUSBLICK.....	- 75 -
7 ZUSAMMENFASSUNG	- 78 -
LITERATURVERZEICHNIS	- 79 -
ABBILDUNGSVERZEICHNIS.....	- 90 -
TABELLENVERZEICHNIS.....	- 91 -
ABKÜRZUNGSVERZEICHNIS	- 92 -
ANHANG	- 94 -
ANTRAGSFORMULAR	- 94 -

STUDIENPROTOKOLL	- 98 -
PROBANDENINFORMATION	- 107 -
COGNITIVE TEST ANXIETY SCALE (CTAS)	- 109 -
COGNITIVE TEST ANXIETY SCALE – DEUTSCHE ÜBERSETZUNG (G-CTAS)	- 111 -
ABBILDUNGEN	- 113 -
TABELLEN	- 117 -
<i>DANKSAGUNG</i>	- 118 -
<i>EIDESSTATTLICHE VERSICHERUNG</i>	- 119 -

Hinweis: Aus Gründen der besseren Lesbarkeit wurde überwiegend die maskuline Form der Anrede verwendet.

1 Einleitung

1.1 Prüfungsangst

1.1.1 Definition und Konstrukt, die „Zwei-Komponenten-Theorie“

„Prüfungsangst stellt eine spezifische Form der Leistungsangst dar, bei der eine Person vor und in einer Prüfungssituation ein Übermaß an Sorge, physiologischer Erregung, mentaler Desorganisation und unkontrollierbaren, selbstwertbedrohlichen Gedanken erlebt, den Anforderungen der Prüfung nicht zu genügen. [...]“ [1]

Prüfungsangst wird zu den Angststörungen gezählt, nach dem „Diagnostic and Statistical Manual of Mental Disorders, 5th Edition“ (DSM-V) [2] wird sie als soziale, nach der „International Statistical Classification of Diseases and Related Health Problems, 10th Revision“ (ICD-10) [3] als spezifische (isolierte) Phobie kodiert [4]. Laut Metzger & Schuster [5] stellt sie gemeinsam mit Lampenfieber und der Angst vor Autoritätspersonen eine Form der Bewertungsangst dar. Eine allgemein akzeptierte und somit verbindliche Definition von Prüfungsangst liegt nicht vor, da sie im Klassifikationssystem ICD-10 nicht explizit als Störung mit Krankheitswert aufgeführt ist [6]. Die meisten Definitionen, so auch die oben genannte von Holodyski, nehmen an, dass es sich bei Prüfungsangst um ein komplexes Konstrukt handelt, welches sich in einer Vielzahl körperlicher Symptome und kognitiver Prozesse manifestieren kann.

Allgemein anerkannt ist, dass sich Prüfungsangst aus zwei Hauptkomponenten, „emotionality“ und „worry“, zusammensetzt (*Zwei-Komponenten-Theorie*) [7]. Selbst wahrgenommene, autonome Reaktionen, die in Prüfungssituationen auftreten, werden der „emotionality“-Komponente zugeschrieben. Hierzu zählen physiologische Phänomene wie Schwitzen, Ansteigen der Herzfrequenz und Nervosität. Sie wird auch als affektive Komponente bezeichnet. Der „worry“-Komponente wird eine Vielzahl an kognitiven Prozessen zugerechnet, ursprünglich von Liebert & Morris beschrieben als „jeder kognitive Ausdruck von Besorgnis über die eigene (Prüfungs-)Leistung“ [7]. Dazu zählen unter anderem der Vergleich der eigenen Leistung mit Vergleichsgruppen, Nachdenken über die Konsequenzen des eigenen Versagens und geringes Selbstbewusstsein sowie Verlust des Selbstwertgefühls. Aufgrund der großen Breite und Komplexität der Prozesse, welche die „worry“-Komponente ausmachen, wird diese auch als kognitive Prüfungsangst bezeichnet.

Während über die Unterscheidung dieser beiden Hauptkomponenten überwiegend Konsens besteht, herrschen unterschiedliche Meinungen, ob deren weitere Differenzierung in Subkomponenten nötig ist. Diese wird unter anderem in Arbeiten von

Sarason [8], Hodapp [9] und Pekrun, Götz et al. [10] favorisiert (*siehe 1.1.7 Messinstrumente von Prüfungsangst*). Auch eine Unterscheidung in aktuell erlebte („state test anxiety“) und habituelle Prüfungsangst („trait test anxiety“) sehen einige Autoren als wegweisend an.

1.1.2 Entstehung

Eine zentrale Rolle bei Überlegungen zur Entstehung von Prüfungsangst nimmt die Frage nach der Interaktion zwischen Prüfungsangst und reduzierter Prüfungsleistung ein. Auf die möglichen Auswirkungen von verminderter Prüfungsleistung soll im Abschnitt *1.1.4.1 Prüfungsangst und Prüfungsleistung* eingegangen werden.

Laut Pekrun & Götz [11] spielt die subjektive Einschätzung bezüglich des eigenen Lernverhaltens, der Prüfungsleistung und der Prüfung selbst eine entscheidende Rolle. Die Erwartung eines Misserfolgs mit damit verbundenen negativen Konsequenzen und der erlebte Kontrollmangel können zur Entstehung von Prüfungsangst führen. Kontrollverlust kann außerdem durch schwierige Übungsaufgaben, erschwerte Lernbedingungen, Erwartungen des Umfelds und wichtige, häufige oder intransparente Prüfungen herbeigeführt werden. Relevant sind also auch Ungewissheiten bezüglich Prüfungsmodalität und -anforderungen.

Kontrolle, beziehungsweise deren Wahrnehmung (engl. „perceived control“), nimmt auch einen zentralen Stellenwert in der Erklärung von Putwain & Aveyard [12] ein. Die Kontrollwahrnehmung konnte als moderierende Variable zwischen Prüfungsangst und Prüfungsleistung identifiziert werden.

Neben Wissensdefiziten und Kontrolle kann auch Selbstbewusstsein eine Rolle im Entstehungsprozess spielen. Reeve, Bonaccio et al. [13] zeigten, dass Selbstbewusstsein bezüglich der eigenen Fähigkeiten zu geringerer erwarteter Prüfungsangst führt, wohingegen Gewissheit über das eigene Versagen das Gegenteil bewirkt. Auch lieferten sie eine Bestätigung für Pekruns Annahmen bezüglich der Prüfung und der Prüfungsumstände selbst: Teilnehmer schätzten keine oder wenige Erfahrungen mit dem Testformat als angstausslösend ein, ebenso Prüfungen, welche mit wichtigen Konsequenzen einhergingen.

1.1.3 Stabilität

Laut den Untersuchungen von Lotz & Sparfeldt [14] bleiben Messwerte für Prüfungsangst über einen langen Zeitraum stabil, nehmen jedoch im Vorfeld einer wichtigen Prüfung zu. Die individuelle kognitive Prüfungsangst (CTAS-Werte, *siehe 1.1.7.3 „Cognitive Test Anxiety Scale“ (CTAS) von Cassady & Johnson*) scheint im Vorfeld von Prüfungen konstant zu bleiben [15]. Nicht ausreichend untersucht sind dagegen Veränderungen der kognitiven Komponente vor, während und nach einer Prüfungssituation.

Die bereits geschilderte Sicht von Angst als relativ stabile Persönlichkeitseigenschaft (engl. „trait“) oder als vorübergehender emotionaler Zustand (engl. „state“) ist bei der Betrachtung der Stabilität ebenfalls wichtig. „State anxiety“ äußert sich durch in einer Situation wahrgenommene Anspannung und Unbehagen, wohingegen „trait anxiety“ die Veranlagung meint, eine Situation als gefährlich oder bedrohlich wahrzunehmen und in diesem Fall mit Symptomen der „state anxiety“ zu reagieren. Spielberger definiert Prüfungsangst als „situationsspezifische Form von trait anxiety“ [16] (zusammengefasst von Spielberger, Anton & Bedell [17]). Sie setzt sich seiner Theorie zufolge aus stabilen und situationsspezifischen Merkmalen zusammen.

1.1.4 Negative Auswirkungen

Die Auswirkungen von Prüfungsangst sind bereits in mehreren Kontexten untersucht worden. Bekannt ist der negative Einfluss auf das subjektive Wohlbefinden von Betroffenen [18, 19] sowie deren Zusammenhang mit sozialen und spezifischen Phobien [20] und deren Korrelation mit depressiven Symptomen [21]. Im Fokus der Forschung stehen vor allem die negativen Auswirkungen auf Prüfungserfolge, auf welche im folgenden Abschnitt eingegangen werden soll.

1.1.4.1 Prüfungsangst und Prüfungsleistung

Bereits früh konnten Zusammenhänge zwischen erhöhter Prüfungsangst und unterdurchschnittlichen Prüfungsleistungen festgestellt werden. Einen guten Überblick zu diesem Thema gibt unter anderem Hembree [22]. Als verantwortlich hierfür wird vor allem die kognitive Komponente angesehen. Während der Einfluss der affektiven Komponente als neutral oder sogar positiv bewertet wurde [8, 17], zeigten sich in Untersuchungen wiederholt Korrelationen zwischen kognitiver Prüfungsangst und geringeren Leistungserwartungen sowie mit geringeren realen Prüfungsleistungen [8, 22-27].

Zentrale Ansätze beim Verständnis dieses Zusammenhangs geben Interferenz- (engl. „interference-hypothesis“) und Defizithypothese (engl. „deficit-hypothesis“). Grundidee der Interferenzhypothese ist, dass Test-irrelevantes Denken und sorgenvolle Gedanken die Aufmerksamkeit von der Prüfungsaufgabe ablenken und so zu geringerer Prüfungsleistung bei prüfungsängstlichen Individuen führen. Einige Forscher sprechen diesbezüglich auch von einer „Angstblockade“ (engl. „anxiety blockage“) [28]. Hier wird also von einem kausalen Zusammenhang der Prüfungsangst auf die Prüfungsleistung ausgegangen. Vertreter dieser Hypothese sind unter anderem Wine [29], Sarason [8] und Hembree [22].

Gegenüber steht dieser Erklärung die Defizithypothese. Die Annahme hierbei ist, dass Betroffene in der Prüfungssituation das Vorliegen eines Wissensdefizits registrieren und auf dieses Erkenntnis mit Prüfungsangstsymptomen reagieren. Der Zusammenhang ist

hier also genau umgekehrt: Geringe Prüfungsleistung führt zur Prüfungsangst. Für diese Hypothese sprechen unter anderem die Forschungsergebnisse von Reeve & Bonaccio [30] und Sommer & Arendasy [31, 32].

Hinweise auf einen umfassenderen Erklärungsansatz sind in Cassadys Arbeiten zur kognitiven Prüfungsangst zu finden. Dieser konnte zeigen, dass kognitive Interferenz bereits in der Prüfungsvorbereitungsphase auftritt [33]. Laut seinen Ergebnissen beeinflusst kognitive Prüfungsangst die Prüfungsvorbereitung, -leistung und -reflexion negativ. Prüfungsängstliche Studierende zeigten nicht nur weniger effektive Studienfertigkeiten als ihre Studienkollegen, auch beschrieben sie im Vorfeld eine negative, bedrohliche Wahrnehmung der bevorstehenden Prüfung und berichteten rückblickend über Gefühle von Hilflosigkeit und Kontrollverlust bezüglich ihrer Prüfungsleistung [34]. Cassady sieht darin die Ursache der Aneignung eines Vermeidungsverhaltens betroffener Studierender in Hinblick auf Lernen und Prüfungsvorbereitung.

Die wechselseitige negative Verstärkung von reduzierter Prüfungsleistung, Vermeidungsverhalten und ineffektiven Vorbereitungsstrategien bietet somit eine weitere mögliche Erklärung für den Zusammenhang zwischen Prüfungsangst und Prüfungsleistung. Auch Pekrun beschreibt einen „circulus vitiosus von steigender Angst und kumulierenden Misserfolgen“ [35].

1.1.5 Reduktion

Einen guten Überblick über neuere Interventionen zur Reduktion von Prüfungsangst geben von der Embse, Barterian et al. [36]. Diese lassen sich mehreren Haupt-Ansatzpunkten zuordnen: behavioristisch, kognitiv, kognitiv-behavioristisch, Aufbau von akademischen Fähigkeiten und Biofeedback. Diese Ansätze haben sich laut der Autoren auch in der Therapie generalisierter Angststörungen als effektiv erwiesen. Vor allem stellten sich Kombinationen aus kompetenzorientierten und behavioristischen oder kognitiven Ansätzen als vielversprechend heraus, ebenso Kombinationen aus Einzel- und Gruppentherapieformaten [37].

Tryon [38] bemängelte, dass zwar aufgrund ihres Zusammenhangs mit Prüfungsleistungen primär die kognitive Komponente behandelt werden sollte, Interventionen jedoch bevorzugt auf die Modifikation der affektiven Komponente abzielen. Pintado und Llamazares [39] berichteten, dass prüfungsängstliche Individuen von beiden Interventionsarten profitieren können. Grundsätzlich seien psychologische Interventionen allerdings nur für Teilnehmer mit irrationaler Angst sinnvoll, für Teilnehmer mit rationaler Angst sei das Trainieren von Lern- und Studienfähigkeiten vielversprechender.

Mehrere Ansätze können also möglicherweise zur Prüfungsangstreduktion führen. Aufgrund der abweichenden Definitionen von Prüfungsangst und demzufolge

unterschiedlichen Messinstrumenten wurde bisher jedoch kein Therapieansatz als Goldstandard akzeptiert.

1.1.5.1 Prüfungssimulationen

Prüfungssimulationen werden vielfältig angewandt, beispielsweise um den Stress einer Prüfung zu simulieren [40] oder beim Sammeln von Prüfungserfahrungen zu unterstützen [41]. Simulationen werden von Teilnehmenden als wichtiges Prüfungsvorbereitungsinstrument akzeptiert [42].

Prüfungssimulationen scheinen aus mehreren Perspektiven geeignet, Prüfungsangst positiv zu beeinflussen. Wie unter 1.1.2 *Entstehung* beschrieben, können verschieden Umstände zu deren Entstehung beitragen, darunter Wissensdefizite, geringes Selbstbewusstsein oder fehlerhafte Selbsteinschätzung des Lernstands und der voraussichtlichen Prüfungsleistung sowie Ungewissheit in Bezug auf Prüfung und Prüfungssituation.

Eine Prüfungssimulation bietet die Möglichkeit, ohne negative Konsequenzen mit Format und Anforderungen der Prüfung vertraut zu machen und gibt einen realistischen Überblick über Lernstand und Prüfungsleistung. Desensibilisierung und Konfrontation mit belastenden Situationen sind zudem wichtiger Bestandteil in der Therapie von Angststörungen [43-45]. Sie werden daher von einigen Autoren [46, 47] auch bei Prüfungsangst empfohlen. Saravanan & Kingston konnten bereits in einer Studie zeigen, dass psychologische Interventionen, welche unter anderem Desensibilisierung nutzten, zu einer Reduktion der Angst bei prüfungsängstlichen Medizinstudierenden führen konnten [48].

Vielversprechend ist die Teilnahme an Prüfungssimulationen auch bezüglich einer verbesserten Prüfungsleistung: Nicht nur eine realistische Einschätzung des Lernstands könnte zu einer Intensivierung der Vorbereitung führen, auch konnte wiederholt gezeigt werden, dass das Testen an sich auch für Medizinstudierende [49] eine effektive Lernmethode darstellt [50, 51].

1.1.6 Geschlechtsspezifische Unterschiede

Wiederholt konnten Studien zeigen, dass Frauen höhere Werte für allgemeine und kognitive Prüfungsangst angeben als Männer [22, 27, 52-55], unabhängig vom kulturellen Hintergrund [56]. Mögliche Erklärungen sind, dass Frauen größere Angst haben als Männer, in Prüfungssituationen zu versagen oder aber dass es Männern häufig schwer fällt zuzugeben, unter Prüfungsangst zu leiden [57].

Erhöhte Prüfungsangst bei Frauen wird nicht zwingend von verminderter Prüfungsleistung begleitet, wie bereits Hembree [22] in einer Metaanalyse feststellen konnte. Jüngere Studien, wie beispielsweise von Núñez-Peña, Suárez-Pellicioni et al. [57] stehen in Einklang mit seinen Ergebnissen. Weibliche Teilnehmer schnitten in

Prüfungen nicht schlechter ab als männliche Teilnehmer, obwohl Frauen mit höherer Prüfungsangst vor Vorträgen und offenen Prüfungsfragen rechneten. Keine Unterschiede zwischen den Geschlechtern zeigten sich dagegen bezüglich Prüfungsangst vor Prüfungen mit Multiple-Choice-Fragen. Die Autoren folgern daraus, dass Frauen über effektive Coping-Strategien für den Umgang mit Prüfungsangst verfügen.

1.1.7 Messinstrumente von Prüfungsangst

Mehrere Fragebögen zur Messung von Prüfungsangst sind bereits etabliert und zum Teil in mehrere Sprachen übersetzt worden. Den Fragebögen liegen unterschiedliche Auffassungen bezüglich der Dimensionalität und Definition von Prüfungsangst zugrunde. Drei weit verbreitete Fragebögen sollen im Folgenden kurz vorgestellt und dabei auf deren Eigenschaften eingegangen werden.

1.1.7.1 „Test Anxiety Inventory“ (TAI) von Spielberger

Unter allen in deutscher Sprache verfügbaren Fragebögen ist das „Prüfungsangstinventar“ (TAI) von Spielberger [58] am weitesten verbreitet, allen voran die überarbeiteten Versionen von Hodapp [9, 59] und Hodapp, Rohrmann et al. [60]. In Hinblick auf die „state- und trait-Theorie“ der Angst (*siehe 1.1.3 Stabilität*), misst das TAI situationsspezifisch die beiden Komponenten „worry“ und „emotionality“. Im Gegensatz dazu besteht das weiterentwickelte „German Test Anxiety Inventar“ (TAI-G) von Hodapp aus vier Subskalen „Aufgeregtheit“ (engl. „emotionality“), „Besorgtheit“ (engl. „worry“), „Interferenz“ (engl. „interference“) und „Mangel an Zuversicht“ (engl. „lack of confidence“) [9, 59, 60]. Alle Items beziehen sich auf die Situation während einer Prüfung. Untersuchungsergebnisse sprechen für eine hohe Konstruktvalidität von TAI-G und seiner Versionen [61, 62]. Die Dimensionen „emotionality“ und „worry“ überschneiden sich in Hinblick auf die Zwei-Komponenten-Theorie jedoch stark [63].

1.1.7.2 „Reactions to Tests“ (RTT) von Sarason

Sarason ist Vertreter der Interferenzhypothese. Somit spielt für ihn die kognitive Interferenz in Hinblick auf das Verständnis von Prüfungsangst eine wichtige Rolle. Er führte eine weitere Differenzierung der von Liebert und Morris beschriebenen zwei in insgesamt vier Komponenten durch: „Worry“ und „cognitive interference“ als Subkomponenten der kognitiven Komponente sowie „tension“ und „bodily symptoms“ als Subkomponenten der affektiven Komponente.

Der von ihm entwickelte Fragebogen „Reactions to Tests“ (RTT) umfasst insgesamt 40 Items mit jeweils 10 Items pro Subskala [8], eine überarbeitete Version [64] besteht aus 20 Items.

Es existieren bereits deutsche Versionen des Fragebogens [65, 66].

1.1.7.3 „Cognitive Test Anxiety Scale“ (CTAS) von Cassady & Johnson

Cassady und Johnson stellten 2002 die „Cognitive Test Anxiety Scale“ (CTAS) vor, welche allein die Ausprägung der kognitiven Komponente erfassen soll.

Sie beinhaltet unterschiedliche Facetten kognitiver Prüfungsangst, darunter „Tendenzen, sich in der Prüfungsvorbereitungsphase und während der Prüfung selbst in Aufgaben-irrelevantes Denken zu verstricken, sich mit anderen zu vergleichen und die Neigung, eindringende oder flüchtende Gedanken während Prüfungen oder in Lernsituationen zu erleben“ [67].

In Pilot-Studien konnte eine hohe Übereinstimmung zu den bereits etablierten und oben vorgestellten Fragebögen RTT und TAI gezeigt werden [67].

CTAS wurde bereits für Studien in den Vereinigten Staaten [15, 34, 67-69], Griechenland [70] und Großbritannien [71] verwendet und ist bereits ins Arabische [72], Chinesische [73, 74] und Spanische [75] übersetzt worden. Es existieren außerdem mehrere überarbeitete Versionen [76-78], diese wurden bereits ins Persische [79] und Türkische [80] übersetzt.

1.2. Medizinstudierende & Medizinstudium

1.2.1 Der Erste Abschnitt der Ärztlichen Prüfung (Physikum)

Ablauf sowie Prüfungen des Medizinstudiums in Deutschland sind in der Approbationsordnung [81] festgelegt. Die Ärztliche Prüfung gliedert sich in drei Teile, wovon der erste Abschnitt der ärztlichen Prüfung (1. Staatsexamen=Physikum) nach dem vorklinischen Studienabschnitt abgehalten wird. Die Regel- und gleichzeitig Mindeststudienzeit hierfür beträgt an der Ludwig-Maximilians-Universität München (LMU) vier Semester [82]. Das Physikum besteht aus einer schriftlichen Aufsichtsarbeit (=schriftliches Physikum) und einem mündlich-praktischen Teil (=mündliches Physikum). Das schriftliche Physikum ist deutschlandweit einheitlich geregelt und findet an zwei Tagen statt. Insgesamt werden 320 Fragen durch das Institut für medizinische und pharmazeutische Prüfungsfragen (IMPP) gestellt. Am ersten Prüfungstag werden 160 Fragen zu den Fächern Physiologie, Physik für Mediziner, Biochemie/Molekularbiologie und Chemie für Mediziner gestellt, am zweiten Tag 160 Fragen zu den Fächern Anatomie, Biologie für Mediziner und medizinische Psychologie & medizinische Soziologie. Das Zeitlimit beträgt vier Stunden pro Tag. Basierend auf der Prozentzahl der richtig beantworteten Fragen wird eine Durchschnittsschulnote aus allen Fächern errechnet.

Im mündlichen Physikum werden die Fächer Anatomie, Biochemie/Molekularbiologie und Physiologie in Kleingruppen zu maximal vier Prüflingen geprüft, davon jeder Prüfling für insgesamt 45 bis 60 min. Es wird eine Gesamtnote errechnet (Schulnoten von 1=sehr

gut bis 5=mangelhaft, nicht bestanden) [81]. Prüfungszeiträume sind deutschlandweit uneinheitlich geregelt, an der LMU findet dieser im Herbst vor und nach dem schriftlichen Physikum statt, im Frühjahr in der Regel ausschließlich davor [83]. Vorläufige Prüfungstermine werden zu Beginn der mündlichen Prüfungsphase veröffentlicht, Prüfer werden mit der Prüfungsladung bekannt geben, welche laut Approbationsordnung spätestens fünf, an der LMU München in der Regel jedoch etwa zehn Tage vor der mündlichen Prüfung versendet wird. Bei Nichtbestehen einer oder beider Teilprüfungen können diese jeweils insgesamt zweimal wiederholt werden.

Die Noten aus schriftlicher und mündlicher Prüfung werden addiert und deren Summe durch zwei geteilt. Die Note wird bis auf die erste Stelle nach dem Komma gerundet. Sie fließt in die Abschlussnote der ärztlichen Prüfung ein und macht ein Drittel der Gesamtnote aus [81].

1.2.2 Medizinstudierende als Risikogruppe

Unter Studierenden ist die Stressbelastung im Vergleich zu anderen Bevölkerungsgruppen in Deutschland hoch [84]. Internationale Studien zeigen nahezu übereinstimmend, dass die Prävalenz von depressiven Symptomen, Angst und das Ausmaß psychischer Belastung unter Studierender der Humanmedizin im Vergleich zu Studierenden anderer Studiengänge oder zur Allgemeinbevölkerung erhöht ist [85-90]. Laut Powell wurden unter Medizinstudierenden außerdem wiederholt hohe Prüfungsangstwerte gemessen [91]. Es scheint sich dabei um anhaltende Phänomene zu handeln: Während die Prävalenz psychischer Belastungen zu Studienbeginn vergleichbar mit der Allgemeinbevölkerung sind, nehmen Depressivität und Ängstlichkeit bereits in den ersten Semestern zu [92], die psychische Gesundheit ab [93]. Diese Entwicklung setzt sich in Abhängigkeit vom Aufbau des Curriculums fort [87] und stellt auch ein relevantes Problem während der späteren Tätigkeit als Arzt dar [94, 95]. Ebenfalls scheint die Burnout-Gefahr von Studienbeginn bis in die Berufstätigkeit anzusteigen [96, 97]. Dies wirkt sich nicht nur ungünstig auf die Lebensqualität der Betroffenen aus, sondern reduziert außerdem die Behandlungsqualität und beeinflusst somit die Patientensicherheit negativ [98].

Neben persönlichen Faktoren wie beispielsweise unsicherem Bindungsstil [99] spielen hohe Arbeitsbelastung [100] und die damit häufig einhergehende soziale Isolation [101] eine wichtige Rolle in diesen Entwicklungen. Auch trägt vermutlich die frühe Konfrontation mit Krankheit und Sterben einen relevanten Teil bei [102, 103].

Studierende assoziieren Stress unter anderem mit Leistungsdruck und Angst vor Überforderung. Die wichtigste hochschulbezogene Ursache für Stress stellen laut Herbst et al. Prüfungen dar [84]. Diese kommen somit außerdem als Auslöser für die geschilderten negativen Phänomene unter Medizinstudierenden in Betracht.

2 Fragestellung und Zielsetzung

Prüfungen stellen eine potentielle Ursache für die erhöhte Stressbelastung und Angstprävalenz unter Medizinstudierenden dar. Unbekannt ist, inwieweit eine relevante Prüfung wie das Physikum Sorgen bei den Studierenden auslösen kann und mit welchen Faktoren diese korreliert sind.

Insbesondere die kognitive Komponente der Prüfungsangst muss aufgrund ihrer negativen Folgen in diesem Zusammenhang untersucht werden. Ein geeignetes Messinstrument für die umfangreiche Erfassung fehlt jedoch, da bisher keine deutschsprachige Version der CTAS existiert. Nicht ausreichend analysiert sind zudem intraindividuelle Stabilität und Verteilung kognitiver Prüfungsangst sowie deren Beeinflussung durch Vorbereitung auf und Teilnahme an einer wichtigen Prüfung.

Ein wenig untersuchtes Format in Hinblick auf die Reduktion von kognitiver Prüfungsangst und Prüfungssorgen sowie auf die Steigerung von Lernaufwand und Prüfungsleistungen sind Prüfungssimulationen. Um deren Eignung als Prüfungsvorbereitungsangebot zu testen, sollten die Zielgruppe des Formats und die Auswirkungen einer Teilnahme untersucht werden.

Hieraus ergeben sich folgende Zielsetzungen für die Arbeit:

- Übersetzung der „Cognitive Test Anxiety Scale“ (CTAS) ins Deutsche und deren psychometrische Validierung an Medizinstudierenden vor dem Physikum
- Verteilungsanalyse der Skalensummenwerte unter Medizinstudierenden unmittelbar vor dem Physikum
- Überprüfung der Stabilität kognitiver Prüfungsangst durch Vergleich von Werten vor und nach dem Physikum
- Bestimmung des Ausmaßes an Prüfungssorgen und Analyse von Zusammenhängen zum besseren Verständnis erhöhter Stressbelastung durch das Physikum
- Durchführung von Prüfungssimulationen des schriftlichen Physikums sowie Untersuchung der Zielgruppe und möglicher Effekte auf kognitive Prüfungsangst, Prüfungssorgen, Lernaufwand und Prüfungsleistungen

3 Material & Methoden

3.1 Aufbau der Studie

3.1.1 Studiendesign

Die Studie wurde zu drei Zeitpunkten im Herbst 2015, Frühjahr 2016 und Herbst 2016 durchgeführt.

Etwa fünf Wochen vor dem schriftlichen Physikum wurde Umfrage I an alle potentiellen Physikumskandidaten (4. vorklinisches Semester oder höher) versendet, Umfrage II nach Bekanntgabe der Prüfungsergebnisse etwa fünf Wochen nach dem schriftlichen Physikum an dieselbe Studierendengruppe. Dazu wurden Anschreiben mit Link zu online-Umfragen via E-Mail versendet. Jeweils drei bzw. vier Wochen vor den schriftlichen Physikumsprüfungen fand eine Examenssimulation (siehe 3.1.1.1 Durchführung von Examenssimulationen) statt. Hierzu wurden dieselben Kohorten via E-Mail eingeladen. Die Teilnahme an Umfragen und Examenssimulationen war freiwillig. Zur Beantwortung der Fragestellungen wurden unterschiedliche Stichproben analysiert und miteinander verglichen. Diese wurden in Abhängigkeit davon gebildet, welche Umfragen beantwortet wurden und ob an der Examenssimulation teilgenommen wurde (Simulationsgruppe) oder nicht (Kontrollgruppe, siehe 3.3 Stichproben). Eine graphische Darstellung des Studiendesigns findet sich in Abbildung 1.

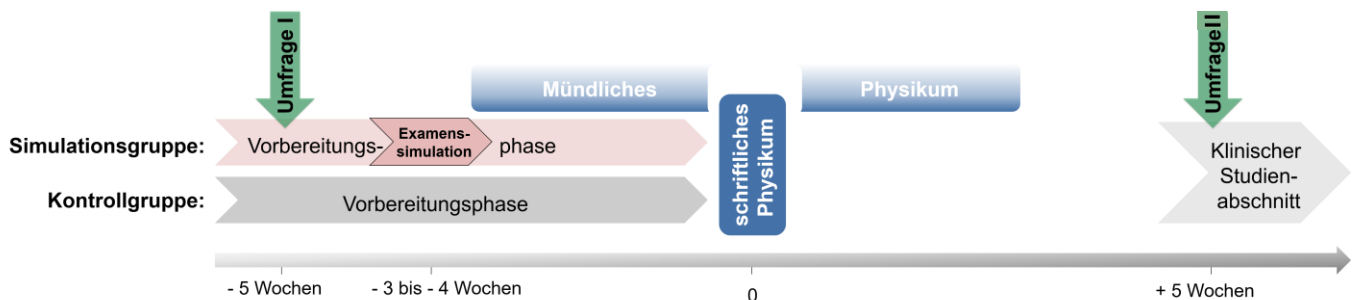


Abbildung 1 - Studiendesign

Simulations- und Kontrollgruppe durchliefen die Vorbereitungsphase auf das Physikum mit bzw. ohne Examenssimulation. Die Umfragen wurden einmal in der Vorbereitungsphase vor der Simulation (Umfrage I) und einmal nach dem Physikum (Umfrage II) an die Studierenden versendet. Während das schriftliche Physikum im Herbst für alle Studierenden zur selben Zeit stattfand, waren Termine für das mündliche Physikum auf die Zeit vor und nach dem schriftlichen Physikum verteilt.

3.1.1.1 Durchführung von Examenssimulationen

Simulationen des schriftlichen Physikums wurden so realitätsgetreu wie möglich durchgeführt. Sie fanden drei (Frühjahr 2016, Herbst 2016) bzw. knapp vier Wochen (Herbst 2015) vor dem realen schriftlichen Physikum statt. Insgesamt nahmen 99 Studierende an den drei durchgeführten Examenssimulationen teil. Zur Teilnahme war die Angabe der Matrikelnummer notwendig.

Die Examenssimulationen fanden in Hörsälen der medizinischen Fakultät der LMU statt. Hilfsmittel wie Mobiltelefon, Fachbücher oder Formelsammlungen waren nicht zugelassen. Analog zum realen Physikum gab es ein Zeitlimit, welches nicht überschritten werden durfte. Während der gesamten Zeit war eine Prüfungsaufsicht anwesend, Studierende, die den Prüfungsablauf oder die Prüfungsatmosphäre störten, wurden verwahrt.

Aufgabenhefte mit kompletten Fragensätzen (insgesamt jeweils 320 Fragen) eines vergangenen Examens (Herbst 2014 bis Herbst 2015) wurden den Teilnehmern zur Bearbeitung bereitgestellt. Dabei wurden jeweils zwei unterschiedliche Fragensätze verteilt, um die Prüfungsbedingungen im Physikum (zwei Prüfungsgruppen mit unterschiedlichen Fragenreihenfolgen) zu imitieren. Layout und Fragenreihenfolge entsprachen den Original-Examensaufgabenheften. Ebenfalls waren die Fragen analog zur Original-Examensprüfung auf die verschiedenen Prüfungstage aufgeteilt.

Die Hefte wurden im Rahmen einer von uns initiierten Kooperation durch den Verlag „Elsevier“ kostenlos zur Verfügung gestellt. Die Hefte enthielten das Logo des Verlags und einen Werbeflyer, der Erhalt war für Veranstalter und Teilnehmer mit keinen Verpflichtungen oder mit der Weitergabe von Daten verbunden.

3.1.2 Datenschutz und Datennutzung

Alle potentiellen Physikumskandidaten, darunter auch die Examenssimulationsteilnehmer, wurden schriftlich mittels Anschreiben und Probandeninformation (*siehe Anhang*) über den Zweck der geplanten Datenerhebung informiert. Die Einwilligung zur Datennutzung erfolgte durch die Aktivierung des Umfragenlinks. Die Teilnahme an den Umfragen und Examenssimulationen waren freiwillig, Aufwandsentschädigungen wurden nicht gezahlt. Allen Teilnehmern stand es frei, die Teilnahme jederzeit und ohne Angabe von Gründen zu widerrufen. Keiner der Teilnehmer machte von dieser Möglichkeit Gebrauch.

Persönliche Daten der Teilnehmer wurden pseudonymisiert mithilfe eines auf Basis der Matrikelnummer generierten Verschlüsselungscodes gespeichert.

3.1.3 Ethikvotum

Ein Antragsformular mit Studienprotokoll (*siehe Anhang*) wurde der Ethikkommission der Medizinischen Fakultät der LMU München im März 2015 vorgelegt und von dieser genehmigt.

3.2 Messinstrumente

Für die Studie wurden zwei Online-Umfragen (Umfrage I und II) erstellt. Diese wurden zu zwei unterschiedlichen Zeitpunkten an die Studierenden versendet (*siehe 3.1.1 Studiendesign*). Beide enthielten eine Skala zur Einschätzung der kognitiven Prüfungsangst (*siehe 3.2.1 Cognitive Test Anxiety Scale (CTAS)*) sowie Fragen nach demographischen Daten (Alter, Geschlecht). In beiden Umfragen war außerdem die Angabe der Matrikelnummer erforderlich. Die Zusatzfragen unterschieden sich in Umfrage I und II voneinander (*siehe 3.2.2 Zusatzmaterial*).

3.2.1 „Cognitive Test Anxiety Scale“ (CTAS)

Wie bereits beschrieben wurde CTAS entwickelt, um ausschließlich die kognitive Komponente der Prüfungsangst zu messen. Für die Untersuchungen wurde der ursprüngliche Fragebogen von Cassady & Johnson [67] verwendet. Dieser besteht aus 27 Items, wovon neun Items invers gepolt sind. Die Beantwortung erfolgt auf einer vierstufigen Likertskala (starke Ablehnung=1 Punkt bis starke Zustimmung=4 Punkte, *vollständiger Fragebogen siehe Anhang*). Nach Umkodierung der invers gepolten Items sind hohe Skalenwerte indikativ für ein hohes Level an kognitiver Prüfungsangst. Untersuchungen haben eine hohe interne Konsistenz (alle gemessenen Werte für Cronbach's alpha >0,88 [75]) und hohe Konstruktvalidität der 27-stufigen Skala gezeigt. Sie erwies sich außerdem als stabiles und konsistentes Messinstrument von kognitiver Prüfungsangst mit einer hohen Vorhersagekraft für Prüfungsleistungen [67].

3.2.1.1 Übersetzung ins Deutsche

Alle 27 Items des Original-Fragebogens (CTAS) wurden mittels Vorwärts-Rückwärts-Verfahren übersetzt. Der Fokus wurde dabei weniger auf eine wörtliche als auf eine inhaltstgetreue Übersetzung gelegt. Eine vorläufige Version wurde mit einer englischen Muttersprachlerin diskutiert. Die dabei erarbeitete Version wurde anschließend von drei deutsch- und englischsprachig aufgewachsenen Personen unabhängig voneinander zurück ins Englische übersetzt. Keiner der Personen war der Original-Fragebogen bekannt. Nachdem Unterschiede zwischen der originalen Version und der Rückübersetzung herausgearbeitet worden waren, einigte man sich auf eine deutsche Version.

3.2.2 Zusatzmaterial

Umfrage I enthielt neben der deutschen Version von CTAS und demographischen Angaben (Alter, Geschlecht) Fragen bezüglich Sorgen vor dem Physikum (Sorgen I), Lernaufwand (Lernaufwand I), erwarteten Prüfungsleistungen und bisherigen Prüfungsleistungen. Sorgen vor dem Physikum und erwartete Prüfungsleistungen wurden getrennt für schriftliches und mündliches Physikum erhoben (*siehe Tabelle 1*).

Umfrage II enthielt ebenfalls die deutsche Version von CTAS und demographische Angaben (Alter, Geschlecht). Zusätzlich enthielt sie Fragen zu retrospektiven Sorgen vor dem Physikum (Sorgen II), retrospektivem Lernaufwand (Lernaufwand II) und realen Prüfungsleistungen. Alle Kategorien wurden getrennt für schriftliches und mündliches Physikum erhoben (*siehe Tabelle 2*).

Kategorie mit verwendeter Frage/ Messinstrument	Antwortskala
Kognitive Prüfungsangst	
Deutsche Übersetzung CTAS (27 Items)	vierstufige Likert-Skala (1=“trifft überhaupt nicht auf mich zu“ bis 4=“trifft vollständig auf mich zu“)
Sorgen vor dem Physikum (Sorgen I)	
Sorgen I (schriftlich): „Das schriftliche Physikum bereitet mir Sorgen“	vierstufige Likert-Skala (1=“trifft nicht zu“ bis 4=“trifft vollständig zu“)
Sorgen I (mündlich): „Das mündliche Physikum bereitet mir Sorgen“	
Lernaufwand (Lernaufwand I)	
„Wie schätzen Sie den Lernaufwand ein, den Sie durchschnittlich für das bevorstehende Physikum aufbringen?“	vierstufige Likert-Skala (1=“sehr niedrig“ bis 4=“sehr hoch“)
Erwartete Prüfungsleistungen	
Erwartete Prüfungsleistungen (schriftlich): „Wie gut werden Sie Ihrer Meinung nach im schriftlichen Physikum abschneiden?“	Schulnoten von 1 bis 5
Erwartete Prüfungsleistungen (mündlich): „Wie gut werden Sie Ihrer Meinung nach im mündlichen Physikum abschneiden?“	
Bisherige Prüfungsleistungen	
Abiturnote	Schulnoten von 1,0 bis 4,0

Tabelle 1 - Inhalt Umfrage I: Kategorie mit zugehörigen Fragen/Messinstrumenten und Antwortskalen

Umfrage I enthielt die deutsche Version von CTAS sowie Fragen zu den Kategorien „Sorgen vor dem Physikum“ (Sorgen I), „Lernaufwand“ (Lernaufwand I), „erwartete Prüfungsleistungen“ und „bisherige Prüfungsleistungen“.

Kategorie mit verwendeter Frage/ Messinstrument	Antwortskala
Kognitive Prüfungsangst	
Deutsche Übersetzung CTAS (27 Items)	vierstufige Likert-Skala (1=“trifft überhaupt nicht auf mich zu“ bis 4=“trifft vollständig auf mich zu“)
Sorgen vor dem Physikum - retrospektiv (Sorgen II)	
Sorgen II (schriftlich): „Das schriftliche Physikum hat mir Sorgen bereitet“	vierstufige Likert-Skala (1=“trifft nicht zu“ bis 4=“trifft vollständig zu“)
Sorgen II (mündlich): „Das mündliche Physikum hat mir Sorgen bereitet“	
Lernaufwand - retrospektiv (Lernaufwand II)	
Lernaufwand II (schriftlich): „Wie schätzen Sie den Lernaufwand ein, den Sie für das schriftliche Physikum aufgebracht haben?“	vierstufige Likert-Skala (1=“sehr niedrig“ bis 4=“sehr hoch“)
Lernaufwand II (mündlich): „Wie schätzen Sie den Lernaufwand ein, den Sie für das mündliche Physikum aufgebracht haben?“	
Reale Prüfungsleistungen	
Reale Prüfungsleistungen (schriftlich): „Mit welcher Note haben Sie im schriftlichen Physikum abgeschnitten?“	Schulnoten von 1 bis 5
Reale Prüfungsleistungen (mündlich): „Mit welcher Note haben Sie im mündlichen Physikum abgeschnitten?“	

Tabelle 2 - Inhalt Umfrage II: Kategorie mit zugehörigen Fragen/Messinstrumenten und Antwortskalen

Umfrage II enthielt die deutsche Version von CTAS sowie Fragen zu den Kategorien „Sorgen vor dem Physikum - retrospektiv“ (Sorgen II), „Lernaufwand - retrospektiv“ (Lernaufwand II) und „reale Prüfungsleistungen“.

3.3 Stichproben

Nur vollständig bearbeitete Umfragen wurden in die Auswertung mit einbezogen. Doppelte Bearbeitungen konnten durch die Angabe der Matrikelnummer identifiziert und ausgeschlossen werden.

Alle Teilnehmer, welche Umfrage I beantwortet hatten, wurden zur Stichprobe 1 gerechnet (Anzahl (N)=339). Stichprobe 1 wurde weiterhin unterteilt in Teilnehmer, welche beide Umfragen beantwortet hatten (Stichprobe 2, N=144) oder ausschließlich Umfrage I (Stichprobe 3, N=195).

Insgesamt hatten 284 Teilnehmer Umfrage II beantwortet. Für die vorliegende Arbeit wurden unter diesen nur jene berücksichtigt, welche bereits an Umfrage I teilgenommen hatten (Stichprobe 2). Antworten von Teilnehmern, welche ausschließlich an Umfrage II teilgenommen hatten, flossen demzufolge in die Auswertungen der vorliegenden Arbeit nicht ein.

Es erfolgte eine weitere Unterteilung der Stichprobe 2 in Simulations- und Kontrollgruppe in Abhängigkeit davon, ob eine Teilnahme an der Examenssimulation erfolgt war (Simulationsgruppe, N=52) oder nicht (Kontrollgruppe, N=92, *siehe 3.1.1 Studiendesign*). Die Stichprobenzuteilung ist graphisch in Abbildung 2 dargestellt.

Detaillierte Beschreibungen der Stichproben bezüglich Umfang, Geschlechterverteilung und Alter finden sich in Tabelle 3.

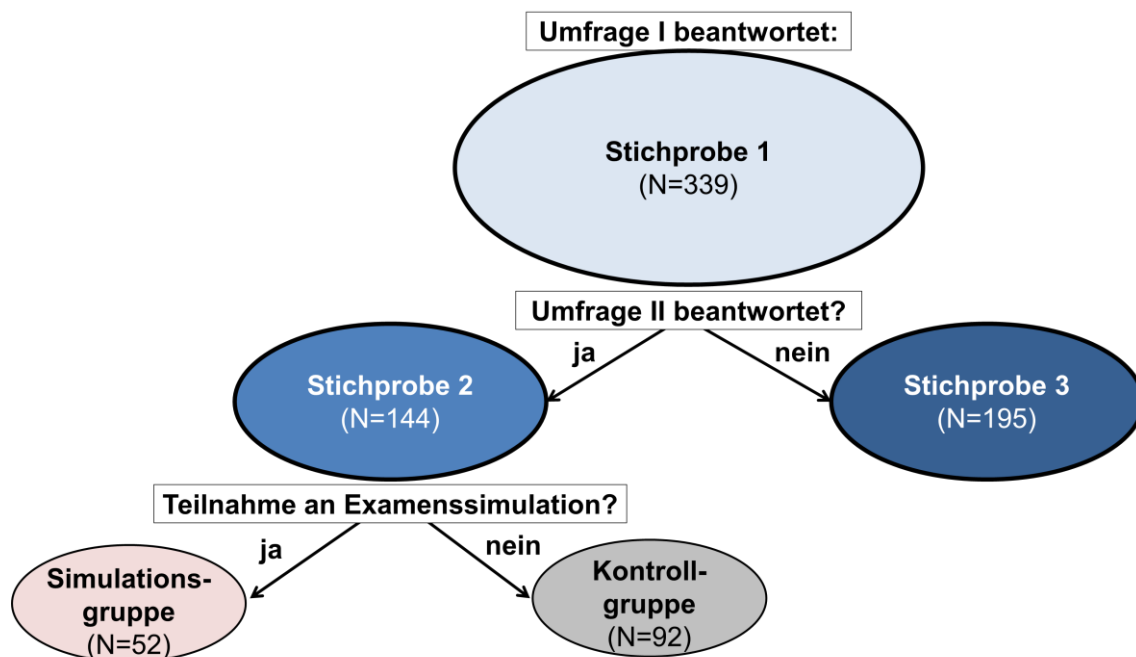


Abbildung 2 - Stichprobenzuteilung

Die Zuteilung zur Stichprobe erfolgte in Abhängigkeit von Beantwortung der Umfragen und Teilnahme an der Examenssimulation.

	Anzahl (N)			Alter (M ± SD)
	gesamt	weiblich	männlich	
Stichprobe 1	339	231 (68,1%)	108 (31,9%)	22,49 ± 4,04
Stichprobe 2	144	90 (62,5%)	54 (37,5%)	21,96 ± 3,15
- Simulationsgruppe	52	30 (57,7%)	22 (42,3%)	21,90 ± 3,29
- Kontrollgruppe	92	60 (65,2%)	32 (34,8%)	21,99 ± 3,09
Stichprobe 3	195	141 (72,3%)	54 (27,7%)	22,89 ± 4,55

Tabelle 3 - Stichprobenbeschreibungen

Dargestellt sind Stichprobenumfänge, Anzahl (N) weiblicher und männlicher Teilnehmer und durchschnittliches Alter (angegeben in Mittelwert (M) mit Standardabweichung (SD)) der verschiedenen Stichproben.

3.4 Statistische Auswertung

Zu Beginn der Analysen wurde eine Codierung der Antworten von CTAS-Items und der Zusatzfragen für alle Stichproben vorgenommen. Codierungen für die einzelnen Fragen sind in Tabelle 1 und 2 dargestellt. Es wurde außerdem eine Umkodierung der neun invers gepolten Items von CTAS durchgeführt und für alle folgenden statistischen Analysen beibehalten.

3.4.1 Instrumente und Programme

Für die statistische Datenauswertung wurden das „Statistical Package for the Social Sciences“ (SPSS) 25 und Microsoft excel 2013 verwendet. Die Parallelanalyse nach Horn [104] wurde mithilfe der webbasierten „Parallel Analysis Engine“ von Patil, Singh et al. [105] durchgeführt. Effektstärken wurden mithilfe des online-Rechners „Berechnung von Effektstärken“ von Lenhard & Lenhard bestimmt [106].

3.4.2 Testverfahren

3.4.2.1 Tests zur Überprüfung von Unterschiedshypothesen

Mittelwerte (M) wurden mittels t-Tests oder nicht-parametrischen Mann-Whitney-Tests bzw. Wilcoxon-Tests für unverbundene oder verbundene Stichproben verglichen und auf Unterschiede überprüft.

Bei ausreichend großen Stichproben ($N > 30$) wurden ohne vorherige Testung des Merkmals auf Normalverteilung t-Tests durchgeführt. Dieses Vorgehen wird unter anderem von Bortz & Schuster [107] empfohlen, da der t-Test bei ausreichend großer Stichprobenanzahl das festgelegte Signifikanzniveau auch einhält, wenn das Merkmal nicht normalverteilt ist. Levene-Tests wurden verwendet, um Varianzgleichheit zu überprüfen.

War die Stichprobengröße klein ($N \leq 30$), wurde das zu überprüfende Merkmal mittels Komogorov-Smirnov-Test auf Normalverteilung überprüft. Lag diese vor, wurden t-Tests wie oben beschrieben durchgeführt. Lag diese nicht vor, wurden der nicht-parametrische Mann-Whitney-Test für unverbundene oder der nicht-parametrische Wilcoxon-Test für verbundene Stichproben verwendet.

Waren die Unterschiede signifikant, wurde Cohens d (d) als Maß für die Effektstärke berechnet. Wie von Cohen [108] empfohlen wurde bei $|d| > 0,2$ die Effektstärke als klein angesehen, bei $|d| > 0,5$ als moderat und bei $|d| > 0,8$ als groß.

3.4.2.2 Korrelationen

Korrelationen wurden mithilfe des Korrelationskoeffizienten nach Pearson (r_p) oder Rangkorrelationskoeffizienten nach Spearman (r_s) berechnet. Analog zum Vorgehen bei der Überprüfung von Unterschiedshypothesen wurde der Korrelationskoeffizient nach Pearson verwendet, wenn die Stichprobengröße ausreichend groß war ($N > 30$) oder eine

Normalverteilung des zu überprüfenden Merkmals vorlag. War die Stichprobe klein ($N \leq 30$) und konnte der Komogorov-Smirnov-Test keine Normalverteilung zeigen, wurde der Rangkorrelationskoeffizient nach Spearman (r_s) verwendet.

Bei $|r| > 0,30$ wurde die Korrelation als moderat angesehen, während sie bei $|r| > 0,50$ als stark betrachtet wurde.

3.4.2.3 Signifikanz

Es wurde ein Signifikanzniveau von $\alpha = 0,05$ festgelegt. P-Werte (p) von $< 0,05$ wurden als statistisch signifikant betrachtet und mit * gekennzeichnet, $p < 0,01$ als statistisch sehr signifikant und mit ** gekennzeichnet, $p < 0,001$ als statistisch hoch signifikant und mit *** gekennzeichnet. Die Nullhypothese wurde anerkannt, wenn $p \geq 0,05$.

3.4.3 Item- und Testanalyse, Validierung von G-CTAS

Item- und Testanalysen der deutschen Version von CTAS wurden an Stichprobe 1 durchgeführt, die Überprüfung der Validität an Stichprobe 2.

3.4.3.1 Deskriptive Statistik und Itemanalyse

Mittelwerte (M) mit Standardabweichung (SD) und Varianz (V) aller Items von CTAS wurden berechnet. „Die psychometrische Itemschwierigkeit ist als Zustimmung zu einem Item in Schlüsselrichtung der Skala definiert“ [109]. Schwierigkeitsindices der einzelnen Items wurden daher über Mittelwerte der Items bestimmt (p_m). Ausgehend von einer vierstufigen Likert-Antwortskala wurde eine optimale Itemschwierigkeit von $p_m = 2,5$ festgelegt (entsprechend einem Schwierigkeitsgrad von 50%) [110]. Die Skala sollte weder zu schwierige noch zu leichte Items enthalten. Dennoch sollten sich die enthaltenen Items in ihrer Schwierigkeit unterscheiden und deren Antworten unterschiedliche Verteilungen aufweisen, um in allen Ausprägungsbereichen inklusive der Randbereiche kognitiver Prüfungsangst differenzieren zu können [109]. Bei auffallend leichten oder schwierigen Items wurden daher auch Varianzen und Histogramme beurteilt. War die Varianz niedrig oder wurde das Antwortformat nicht voll ausgeschöpft, wurde das Item auf problematische inhaltliche Formulierungen überprüft und dessen Ausschluss diskutiert.

Eigentrennschärfen der Items wurden mittels Produkt-Moment-Korrelation zwischen dem Item und dem um das Item bereinigten Gesamtscore der Skala (=part-whole-korrigierter Trennschärfeindex (r_{it})) berechnet. Items mit relativ niedrigen Trennschärfeindices wurden bezüglich ihres Inhalts überprüft und falls nötig entfernt. Nach Entfernung eines Items wurde erneut eine Berechnung der Trennschärfen der verbliebenen Items vorgenommen.

3.4.3.2 Exploratorische Faktorenanalyse (EFA)

Die Faktorenstruktur der Skala wurde mittels exploratorischer Faktorenanalyse (EFA) analysiert und die Ergebnisse mit bisherigen Untersuchungen verglichen. Als Kriterien zur Bewertung der Durchführbarkeit der EFA wurden der Kaiser-Meyer-Olkin-Koeffizient (KMO) und Measure of Sample Adequacy-Koeffizienten (MSA-Koeffizienten) der einzelnen Items errechnet sowie der Bartlett-Test durchgeführt. Zur Bestimmung der Faktorenanzahl wurden im Rahmen einer Hauptkomponentenanalyse Kaiser's Eigenwertkriterium größer eins [111] und der Scree-Plot nach Cattell [112] beurteilt sowie eine Parallelanalyse nach Horn [104] durchgeführt. Als Methode für die EFA wurde eine Hauptachsenanalyse (PAF) mit obliquen Promax-Rotation gewählt.

Geringe Ladungen der Items auf einen Faktor ($<0,3$) wurden in den Analysen nicht berücksichtigt. Items wurden einem Faktor zugeordnet, wenn sie allein auf diesen Faktor luden. Bei Ladungen auf mehrere Faktoren (Querladungen) wurde das Item entfernt. Ausnahme bildeten Querladungen mit relativ hohen Ladungen auf nur einen Faktor (Ladungsunterschied $>0,3$), in diesem Fall wurde das Item diesem Faktor zugeordnet.

3.4.3.3 Skalenreliabilität

Zur Bestimmung der Reliabilität der durch Itemanalyse und EFA gebildeten Skala (G-CTAS) wurde deren interne Konsistenz mittels Cronbach's alpha berechnet. Bei Cronbach's alpha $>0,8$ galt die Skala als reliabel [113].

3.4.3.4 Inhaltsvalidität

Zur Überprüfung der Inhaltsvalidität wurden Korrelationen zwischen G-CTAS-Werten in Umfrage I (G-CTAS I) und G-CTAS-Werten in Umfrage II (G-CTAS II) mit erwarteten Prüfungsleistungen im Physikum, realen Prüfungsleistungen im Physikum und bisherigen Prüfungsleistungen (Abiturnote) berechnet. Aufgrund bisheriger Studienergebnisse bezüglich des Zusammenhangs zwischen kognitiver Prüfungsangst und Prüfungsleistung (siehe 1.1.4.1 *Prüfungsangst und Prüfungsleistung*) wurden positive Korrelationen erwartet.

3.4.4 Verteilungsanalyse kognitiver Prüfungsangst

Um die Verteilung kognitiver Prüfungsangst unter Medizinstudierenden vor dem Physikum zu untersuchen, wurden G-CTAS-Werte (G-CTAS I) von Stichprobe 1 errechnet und deren Verteilung ermittelt (Mittelwert (M), Standardabweichung (SD), Maximum (Max.), Minimum (Min.), Modus, Median).

3.4.4.1 Geschlechtsunterschiede

Die oben beschriebenen Verteilungsanalysen wurden zusätzlich getrennt nach Geschlechtern vorgenommen. Um Unterschiede zwischen den Geschlechtern zu untersuchen, wurden G-CTAS-Werte (G-CTAS I) von weiblichen und männlichen

Teilnehmern sowie deren Verteilungen miteinander verglichen. Aufgrund bisheriger Studienergebnisse wurden höhere Werte bei den weiblichen Teilnehmern erwartet (siehe 1.1.6 Geschlechtsspezifische Unterschiede).

3.4.5 Selektionsbias

Vor Durchführung der folgenden Analysen wurde das Vorliegen eines Selektionsbias zwischen Teilnehmern, welche ausschließlich Umfrage I (Stichprobe 3) und Teilnehmern, welche Umfrage I und II beantwortet hatten (Stichprobe 2) untersucht. Stichprobe 2 und 3 wurden hierzu auf Unterschiede bezüglich Alter, Geschlecht, kognitiver Prüfungsangst (G-CTAS I), Sorgen vor dem Physikum (Sorgen I), Lernaufwand (Lernaufwand I), erwarteten Prüfungsleistungen und Abiturnoten miteinander verglichen.

3.4.6 Stabilität kognitiver Prüfungsangst

Die Überprüfung der Stabilität kognitiver Prüfungsangst wurde an der Kontrollgruppe vorgenommen. Dazu wurden mithilfe verbundener Tests Unterschiede kognitiver Prüfungsangst vor (G-CTAS I) und nach dem Physikum (G-CTAS II) untersucht und die Werte miteinander korreliert. Um geschlechtsspezifische Unterschiede bezüglich der Stabilität ausschließen zu können, wurden die Untersuchungen außerdem getrennt nach Geschlechtern durchgeführt.

3.4.7 Ausmaß von Sorgen vor dem Physikum und mögliche Zusammenhänge

Die Analysen wurden an Stichprobe 2 durchgeführt.

Zu Beginn wurde das Ausmaß an Sorgen vor dem schriftlichen und mündlichen Physikum in Stichprobe 2 analysiert. Durch Vergleich von Mittelwerten wurde untersucht, welche Prüfungsmodalität den Studierenden im Vorfeld der Prüfungen größere Sorgen bereitete. Mittelwerte von Sorgen vor schriftlichen und mündlichen Physikum wurden außerdem zwischen den Geschlechtern verglichen.

Um mögliche Zusammenhänge zu untersuchen, wurden kognitive Prüfungsangst (G-CTAS I), Lernaufwand (Lernaufwand I), erwartete Prüfungsleistungen, reale Prüfungsleistungen und Abiturnoten mit Sorgen vor dem schriftlichen (Sorgen I – schriftlich) und mündlichen Physikum (Sorgen I – mündlich) korreliert. Die Analysen wurden außerdem getrennt nach Geschlechtern durchgeführt.

3.4.8 Examenssimulation

Effekte der Examenssimulation wurden durch Vergleich von Simulations- und Kontrollgruppe durchgeführt.

3.4.8.1 Zielgruppe der Examenssimulation

Um Rückschlüsse auf die Zielgruppe der Examenssimulation ziehen zu können, wurden Simulations- und Kontrollgruppe auf Unterschiede bezüglich Alter, Geschlecht,

kognitiver Prüfungsangst (G-CTAS I), Sorgen vor dem Physikum (Sorgen I), Lernaufwand (Lernaufwand I), erwarteten Prüfungsleistungen und bisheriger Prüfungsleistung (Abiturnote) verglichen.

3.4.8.2 Effekte der Examenssimulation

Untersucht wurden mögliche Effekte der Examenssimulation auf kognitive Prüfungsangst, Sorgen vor dem Physikum, Lernaufwand und reale Prüfungsleistungen. Mittelwerte der kognitiven Prüfungsangst in Umfrage II (G-CTAS II), der Sorgen vor dem schriftlichen und mündlichen Physikum (Sorgen II), des retrospektiven Lernaufwands auf das schriftliche und mündliche Physikum (Lernaufwand II) und der realen Prüfungsleistungen im schriftlichen und mündlichen Physikum wurden zwischen Simulations- und Kontrollgruppe verglichen.

Außerdem wurden für beide Gruppen individuelle Unterschiede von kognitiver Prüfungsangst (G-CTAS I und II) und Sorgen vor dem Physikum im mündlichen und schriftlichen Physikum (Sorgen I und II) sowie individuelle Unterschiede zwischen erwarteten Prüfungsleistungen und realen Prüfungsleistungen im schriftlichen und mündlichen Physikum mittels Tests für verbundene Stichproben untersucht.

Alle Untersuchungen wurden zusätzlich zur Aufdeckung eventueller geschlechtsspezifischer Unterschiede getrennt nach Geschlechtern durchgeführt.

4 Ergebnisse

4.1 Item- und Testanalyse, Validierung von G-CTAS

4.1.1 Deskriptive Statistik und Itemanalyse

Deskriptive Statistiken der Items (M, SD, V) sind in Tabelle 4 dargestellt.

Item	M	SD	V
1. Bevorstehende Prüfungen bereiten mir schlaflose Nächte.	2,07	± 0,83	0,69
2. Während einer wichtigen Prüfung denke ich darüber nach, ob die anderen Studenten vielleicht besser abschneiden als ich.	2,16	± 1,02	1,04
3. Mir fällt es weniger schwer als dem durchschnittlichen Studenten, Prüfungsanweisungen präzise aufzufassen. ¹	2,46	± 0,82	0,66
4. Bei wichtigen Klausuren oder Intelligenztests tendiere ich dazu, plötzlich wie gelähmt zu sein.	1,65	± 0,79	0,62
5. Ich bin vor Tests weniger aufgeregt als der durchschnittliche Student. ¹	2,83	± 0,95	0,89
6. Während einer Prüfung denke ich darüber nach, welche Konsequenzen es hätte, wenn ich versage.	2,11	± 1,03	1,06
7. Zu Beginn einer Prüfung bin ich so nervös, dass ich oft nicht klar denken kann.	2,15	± 1,02	1,04
8. Die Aussicht, in einem meiner Kurse einen Test schreiben zu müssen, bereitet mir keine Sorgen. ¹	2,65	± 0,97	0,94
9. Ich bin ruhiger in Testsituationen als der durchschnittliche Student. ¹	2,73	± 0,88	0,77
10. Ich habe weniger Schwierigkeiten als der durchschnittliche Student, Lehrbuchinhalte zu lernen. ¹	2,71	± 0,84	0,70
11. Ich bekomme einen Blackout, wenn ich mich unter Druck für eine Antwort im Test entscheiden muss.	1,71	± 0,83	0,69
12. Während einer Prüfung kommt mir öfter der Gedanke, dass ich vielleicht nicht sehr klug sein könnte.	1,81	± 0,98	0,95
13. Mir fallen Tests unter Zeitdruck besonders leicht. ¹	3,11	± 0,78	0,61
14. Während einer Klausur werde ich so nervös, dass ich Dinge vergesse, die ich eigentlich weiß.	2,17	± 0,91	0,83

15. Nach einer Prüfung habe ich das Gefühl, dass ich es eigentlich hätte besser machen können.	2,58	± 0,96	0,93
16. Ich mache mir größere Gedanken darüber, in Tests gut abzuschneiden, als ich eigentlich sollte.	2,49	± 1,03	1,06
17. Bevor ich zu einer Prüfung antrete, fühle ich mich selbstsicher und entspannt. ¹	3,14	± 0,82	0,67
18. Während ich an einer Prüfung teilnehme, fühle ich mich selbstsicher und entspannt. ¹	2,76	± 0,83	0,69
19. Während einer Prüfung habe ich das Gefühl, dass ich mich nicht besonders gut schlage.	2,15	± 0,75	0,56
20. Wenn ich an einer anspruchsvollen Prüfung teilnehme, fühle ich mich schon als Versager, bevor ich überhaupt angefangen habe.	1,63	± 0,87	0,76
21. Wenn ich im Test auf eine unerwartete Frage stoße, fühle ich mich eher herausgefordert, als panisch. ¹	2,75	± 0,88	0,77
22. Ich bin kein Mensch für Prüfungen, weil meine Prüfungsleistung nicht zeigt, wie viel ich eigentlich über das Thema weiß.	2,17	± 0,91	0,83
23. Ich bin nicht gut im Schreiben von Prüfungen.	1,82	± 0,82	0,68
24. Wenn ich mein Prüfungsexemplar bekomme, dauert es erstmal eine Weile, bis ich mich soweit beruhigt habe, dass ich klar denken kann.	1,49	± 0,80	0,65
25. Ich fühle mich ziemlich unter Druck, gute Noten zu schreiben.	2,33	± 1,06	1,12
26. Ich schneide nicht besonders gut in Prüfungen ab.	2,09	± 0,89	0,79
27. Ich mache oft Leichtsinnfehler in Prüfungen, weil ich so nervös bin.	2,36	± 0,89	0,79

Tabelle 4 - Deskriptive Statistik der deutschen CTAS-Items

Stichprobe 1, N=339. Dargestellt sind die 27 Items der deutschen Übersetzungen von CTAS mit Mittelwert (M), Standardabweichung (SD) und Varianz (V). Eine Umkodierung aller invers gepolten Items (mit ¹ gekennzeichnet) war vor den statistischen Analysen vorgenommen worden.

Die Schwierigkeit der Items lag zwischen $p_m=1,49$ (Item 24) und $p_m=3,14$ (Item 17). Die durchschnittliche Itemschwierigkeit (p_s) betrug 2,30. Itemschwierigkeiten sind graphisch in Abbildung 3 dargestellt.

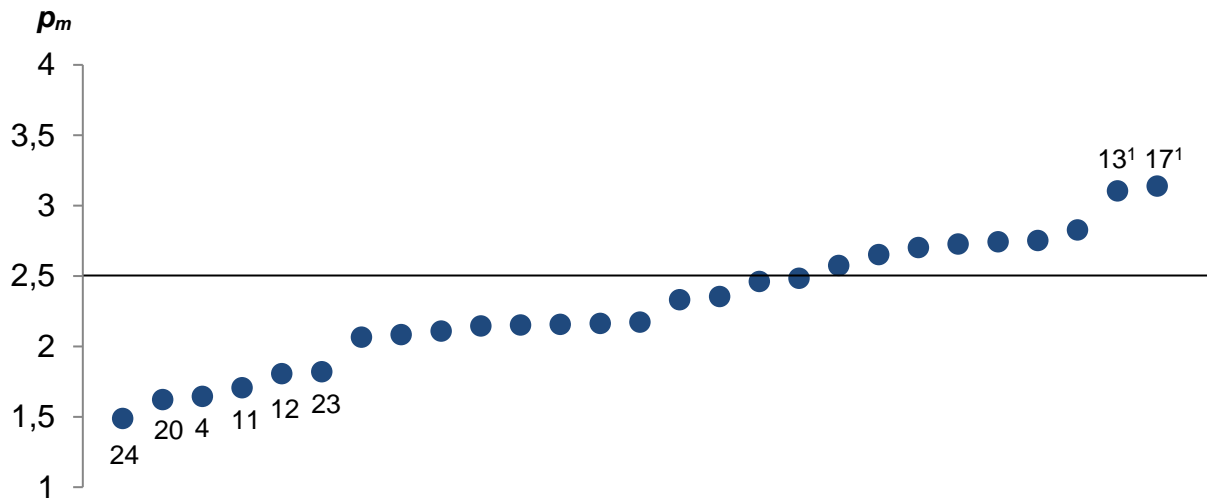


Abbildung 3 - Itemschwierigkeiten

Stichprobe 1, $N=339$. Schwierigkeiten (p_m) der 27 Items sind in aufsteigender Größe abgebildet. Invers gepolte Items waren zu Beginn der Analysen umkodiert worden und sind mit ¹ gekennzeichnet. Die Gerade bei $p_m=2,5$ markiert den optimalen Schwierigkeitsgrad eines Items bei Verwendung einer vierstufigen Likert-Skala.

Histogramme auffallend schwieriger Items (4, 11, 12, 20, 23, 24) und auffallend leichter Items (13, 17) sind in im Anhang abgebildet (siehe Abbildungen 13–20). Item 4, 11, 12, 20, 23 und 24 zeigten linksschiefe, Item 13 und 17 rechtsschiefe Verteilungen. Das volle Antwortformat wurde bei allen Items ausgeschöpft. Varianzen aller 27 Items lagen zwischen $V=0,56$ (Item 19) und $V=1,12$ (Item 25). Item 4, 11, 13, 17, 23 und 24 wurden inhaltlich überprüft, da sie auffallend hohe bzw. niedrige Schwierigkeiten zeigten und ihre Varianzen außerdem im unteren Drittel lagen. Als problematisch konnte die Formulierung von Item 13 identifiziert werden. Item 13 wurde daraufhin von der Skala entfernt, während die übrigen Items als unproblematisch bezüglich ihrer Formulierung eingestuft und für die folgenden Analysen beibehalten wurden.

Anfängliche Trennschärfeindices der Items sind tabellarisch in Tabelle 5 und graphisch in Abbildung 4 dargestellt.

Item	r_{it}
1.	0,57
2.	0,43
3. ¹	0,40
4.	0,64
5. ¹	0,52
6.	0,59
7.	0,67
8. ¹	0,58
9. ¹	0,62
10. ¹	0,36
11.	0,57
12.	0,61
14.	0,63
15.	0,59
16.	0,38
17. ¹	0,59
18. ¹	0,61
19.	0,61
20.	0,68
21. ¹	0,55
22.	0,52
23.	0,58
24.	0,54
25.	0,26
26.	0,43
27.	0,59

Tabelle 5 - Trennschärfeindices

Stichprobe 1, $N=339$. Dargestellt sind Part-whole korrigierte Diskriminationsindices (r_{it}) der verbliebenen 26 Items. Invers gepolte Items waren zu Beginn der Analysen umkodiert worden und sind mit ¹ gekennzeichnet.

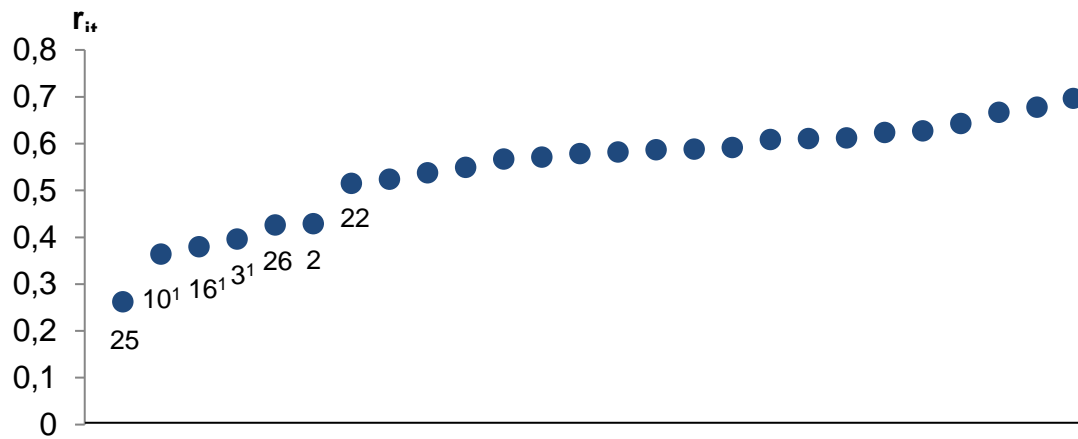


Abbildung 4 - Trennschärfeindices

Stichprobe 1, N=339. Part-whole korrigierte Diskriminationsindices (r_{it}) der verbliebenen 26 Items zu Beginn der Berechnung sind in aufsteigender Größe abgebildet. Invers gepolte Items waren zu Beginn der Analysen umkodiert worden und sind mit ¹ gekennzeichnet.

Die geringste Trennschärfe wurde für Item 25 gemessen ($r_{it}=0,26$). Eine inhaltliche Überprüfung des Items ergab eine zweifelhafte Passung zur restlichen Skala. Ebenfalls konnte dies bei Item 3, 10 und 26 beobachtet werden, worauf diese vier Items von der Skala entfernt wurden.

Item 2 und 16 wurden trotz relativ niedrigen Trennschärfeindices vorläufig nicht von der Skala entfernt, da die inhaltliche Passung beider Items als ausreichend eingeschätzt wurde.

Nach Entfernung eines Items wurde wie beschrieben (*siehe 3.4.3.1 Deskriptive Statistik und Itemanalyse*) eine erneute Berechnung der Trennschärfeindices vorgenommen. Der ausführliche Berechnungsprozess ist tabellarisch im Anhang abgebildet (*siehe Tabelle 18*).

4.1.2 Exploratorische Faktorenanalyse (EFA)

Alle nötigen Voraussetzungen zur Durchführung einer EFA wurden erfüllt (KMO=0,94, alle MSA-Koeffizienten >0,87 (Item 16); Bartlett-Test auf Sphärität: ungefähres Chi-Quadrat (χ^2)=3238,26 bei 231 Freiheitsgraden, Signifikanz nach Bartlett <0,001).

Kaiser's Kriterium zeigte eine Faktorenzahl von vier an, welche gemeinsam 57,5% der Varianz aufklärten. Die Betrachtung des Scree-Plots (siehe Abbildung 5) ergab eine Anzahl von zwei oder vier Faktoren, die Parallel-Analyse eine Faktorenanzahl von zwei. Da die Anzahl der Faktoren bei der Betrachtung von Eigenwerten häufig überschätzt wird [114] und die beiden ersten Faktoren allein 46,6% der Varianz aufklärten (Faktor 1: 39,4%, Faktor 2: 7,2%) wurden zwei Faktoren beibehalten und eine PAF mit Promax-Rotation durchgeführt.

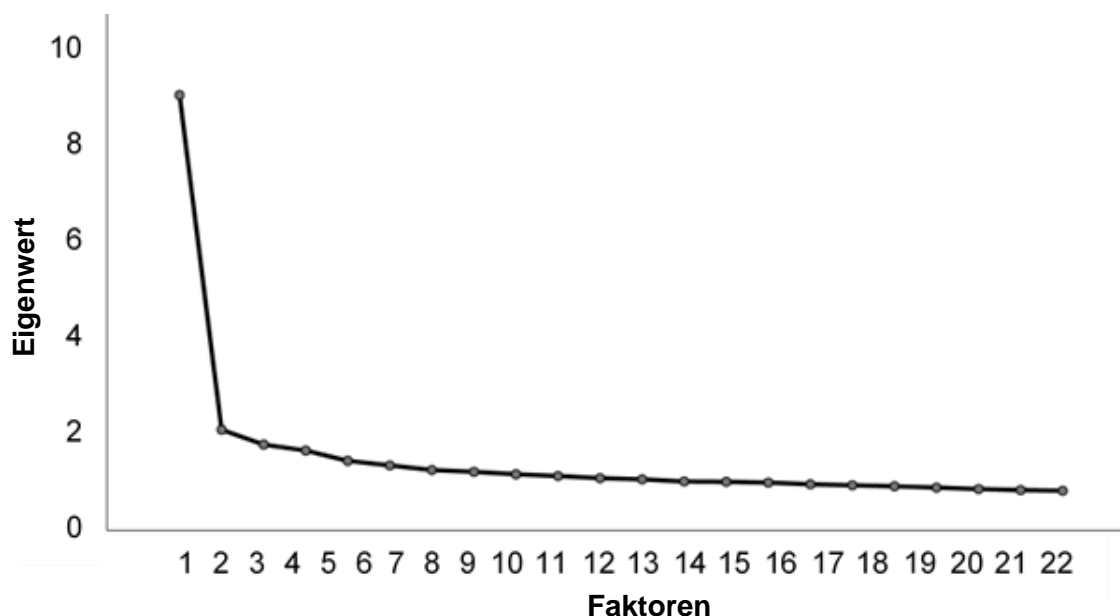


Abbildung 5 – Scree-Plot

Stichprobe 1, N=339. Beurteilt wird ein bedeutender Eigenwertabfall, sichtbar durch einen „Knick“ im Eigenwertverlauf. Gezählt werden die Eigenwerte vor dem Knick.

Ergebnisse der Faktorenanalyse sind in Tabelle 6 dargestellt. Alle invers gepolten Items luden auf Faktor 2. Item 1 und 7 zeigten Querladungen, während Item 2 und 16 auf keinen der beiden Faktoren ausreichend luden.

Item	Faktor 1	Faktor 2
1	0,35	0,31
2		
4	0,62	
5 ¹		0,79
6	0,52	
7	0,39	0,40
8 ¹		0,58
9 ¹		0,87
11	0,53	
12	0,48	
14	0,70	
15	0,65	
16		
17. ¹		0,80
18 ¹		0,56
19	0,45	
20	0,49	
21 ¹		0,45
22	0,79	
23	0,78	
24	0,41	
27	0,65	

Tabelle 6 - Exploratorische Faktorenanalyse

Stichprobe 1, N=339. Eine Exploratorische Faktorenanalyse (EFA) wurde mit 22 der ursprünglich 27 Items durchgeführt. Als Methode wurde eine Hauptachsenanalyse (PAF) mit Promax-Rotation verwendet. Items mit alleinigen, hohen Ladungen auf Faktor 1 (=G-CTAS) sind fett markiert. Invers gepolte Items waren zu Beginn der Analysen umkodiert worden und sind mit ¹ gekennzeichnet.

Item 1, 2, 7 und 16 wurden von der Skala entfernt. Faktor 1 bestand somit aus allen verbliebenen in Schlüsselrichtung gepolten Items, Faktor 2 aus allen verbliebenen invers gepolten Items. Da die Korrelation zwischen den beiden Faktoren nur 0,69 betrug, wurden auch alle Items des zweiten Faktors von der Skala entfernt. Die so generierte eindimensionale Skala (G-CTAS) umfasste zwölf Items, ihre durchschnittliche Schwierigkeit lag bei $p_s=1,97$. G-CTAS-Werte korrelierten mit Summenwerten der 27-Item-Skala stark ($r_p=0,93$) und hoch signifikant ($p<0,001$).

4.1.3 Skalenreliabilität

Cronbach's alpha von G-CTAS war 0,89.

4.1.4 Inhaltsvalidität

Die Überprüfung der Inhaltsvalidität wurde an Stichprobe 2 durchgeführt. Korrelationen (r_p) zwischen G-CTAS-Werten (G-CTAS I und II) mit Abiturnoten sowie mit erwarteten und realen Prüfungsleistungen im mündlichen und schriftlichen Physikum sind in Tabelle 7 dargestellt.

Mit Ausnahme der Korrelation zwischen realer Prüfungsleistung im mündlichen Physikum und G-CTAS I waren alle Korrelationen zwischen G-CTAS-Werten und erwarteten sowie realen Prüfungsleistungen im Physikum moderat und sehr signifikant. G-CTAS II korrelierte mit erwarteter Prüfungsleistung im schriftlichen Physikum und realen Prüfungsleistungen im schriftlichen und mündlichen Physikum stärker als G-CTAS I. Abiturnoten korrelierten nicht mit G-CTAS I oder II.

	G-CTAS I	G-CTAS II
Erwartete Prüfungsleistungen im Physikum (N=144):		
- schriftlich	0,32**	0,37**
- mündlich	0,43**	0,39**
Reale Prüfungsleistungen im Physikum (N=135):		
- schriftlich	0,37**	0,44**
- mündlich	0,26**	0,39**
Abiturnote (N=140):	0,18*	0,14

Tabelle 7 - Überprüfung der Inhaltsvalidität: Korrelationen zwischen G-CTAS-Werten mit erwarteten und realen Prüfungsleistungen

Stichprobe 2, N=144. G-CTAS-Werte in Umfrage I (G-CTAS I) und II (G-CTAS II) wurden mit erwarteten und realen Prüfungsleistungen im schriftlichen und mündlichen Physikum sowie mit Abiturnoten korreliert (r_p). Da es sich um freiwillige Angaben handelte, liegen Daten teilweise nicht für die gesamte Stichprobe vor. Signifikante Unterschiede sind mit * ($p < 0,05$), ** ($p < 0,01$) bzw. *** ($p < 0,001$) gekennzeichnet.

4.2 Verteilungsanalyse kognitiver Prüfungsangst

Deskriptive Statistiken (M, SD, Min., Max., Modus, Median) von G-CTAS I der Stichprobe 1 sind in Tabelle 8 aufgeführt. Alle Werte sind zusätzlich getrennt nach Geschlechtern angegeben. Eine graphische Darstellung der G-CTAS I-Verteilungen für Stichprobe 1 findet sich in Abbildung 6. Der Komogorov-Smirnov-Test lehnte eine Normalverteilung ab ($p < 0,001$).

	M	SD	Min.	Max.	Modus	Median
Stichprobe 1 (N=339)	23,64	7,16	12	46	21	22,00
weiblich (N=231)	24,32	6,91	12	43	24	24,00
männlich (N=108)	22,19	7,50	13	46	18	20,00

Tabelle 8 - Deskriptive Statistiken von G-CTAS-Werten

Angegeben sind Mittelwerte (M), Standardabweichung (SD), Minimum (Min.), Maximum (Max.), Modus und Median. Alle Werte sind außerdem getrennt für weibliche und männliche Teilnehmer angegeben.

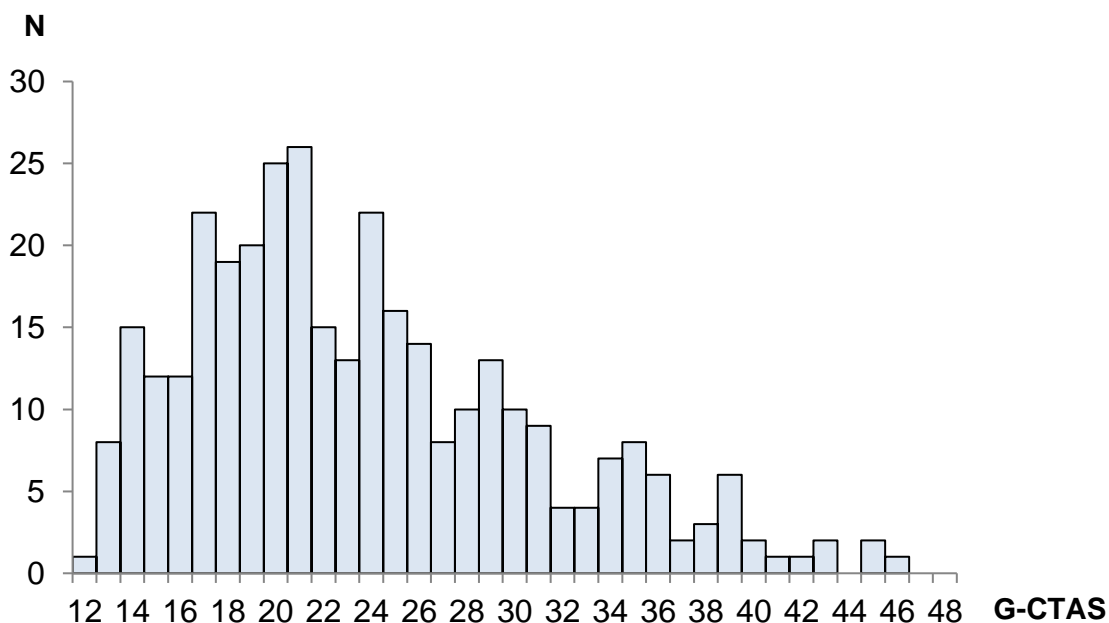


Abbildung 6 - Verteilung von G-CTAS-Werten

Im Histogramm sind G-CTAS-Werte der Stichprobe 1 (N=339) dargestellt.

4.2.1 Geschlechtsunterschiede

Histogramme von G-CTAS I getrennt nach Geschlechtern sind in Abbildungen 7 und 8 dargestellt. Deskriptive Statistiken getrennt nach Geschlechtern finden sich in Tabelle 8. G-CTAS I unterschied sich zwischen weiblichen und männlichen Teilnehmern signifikant ($p < 0,05$; mittlere Differenz: 2,12, T: 2,56, df: 337). Die Stärke des Effekts war klein ($d = 0,3$).

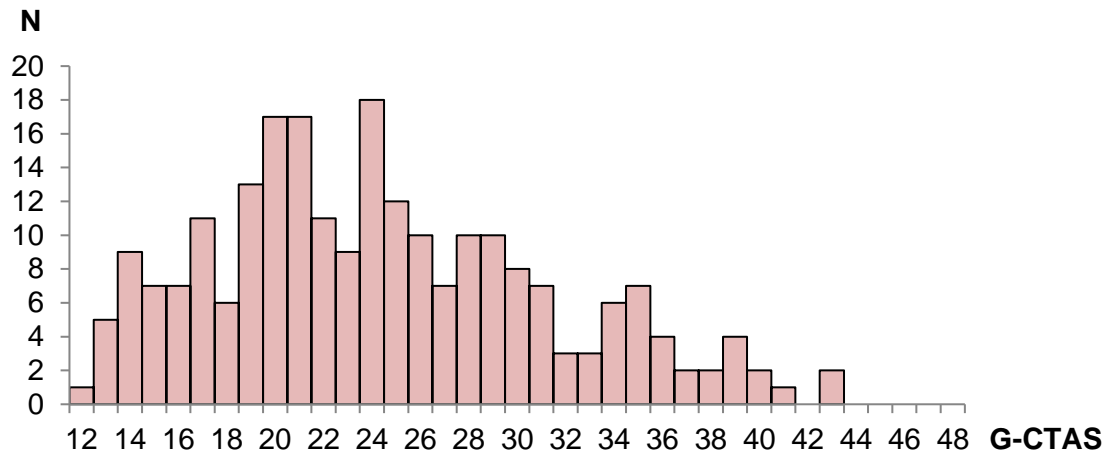


Abbildung 7 – Verteilung von G-CTAS-Werten weiblicher Teilnehmer

Im Histogramm sind G-CTAS-Werte der weiblichen Teilnehmer der Stichprobe 1 (N=231) dargestellt.

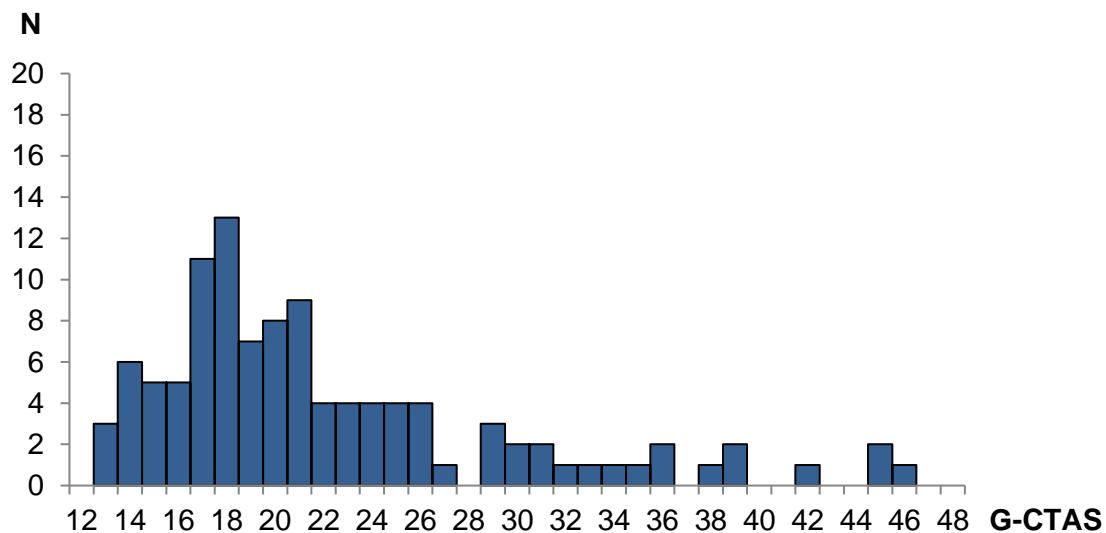


Abbildung 8 – Verteilung von G-CTAS-Werten männlicher Teilnehmer

Im Histogramm sind G-CTAS-Werte der männlichen Teilnehmer der Stichprobe 1 (N=108) dargestellt.

4.3 Selektionsbias

Zwischen Stichprobe 2 und 3 bestanden signifikante Unterschiede bezüglich Alter ($p < 0,036$; mittlere Differenz: -0,929, T: -2,105, df: 337) und kognitiver Prüfungsangst (G-CTAS I; $p < 0,042$; mittlere Differenz: -1,596, T: -2,039, df: 337). Die Effektstärken waren klein ($d(\text{Alter}) = 0,232$; $d(\text{G-CTAS I}) = 0,225$). Keine signifikanten Unterschiede konnten bezüglich Sorgen vor dem Physikum (Sorgen I), Lernaufwand (Lernaufwand I), erwarteten Prüfungsleistungen, Abiturnoten oder der Geschlechterverteilung festgestellt werden (*siehe Tabelle 9*).

	Stichprobe 2 (N=144)	Stichprobe 3 (N=195)	
	M ± SD	M ± SD	Δ
Alter	21,96 ± 3,15	22,89 ± 4,55	0,93*
G-CTAS I	22,72 ± 6,68	24,32 ± 7,43	1,60*
Sorgen I			
- schriftliches Physikum	2,10 ± 0,90	2,29 ± 0,89	0,19
- mündliches Physikum	3,49 ± 0,73	3,51 ± 0,71	0,02
Lernaufwand I	3,23 ± 0,74	3,31 ± 0,70	0,08
Erwartete Prüfungsleistungen			
- schriftliches Physikum	2,43 ± 0,62	2,53 ± 0,75	0,10
- mündliches Physikum	2,85 ± 0,82	2,94 ± 0,87	0,09
Abiturnote	1,50 ± 0,51 (N=140)	1,60 ± 0,59 (N=191)	0,10
Geschlechterverteilung (prozentual)			
- weiblich	62,5%	72,3%	9,8%
- männlich	37,5%	27,7%	9,8%

Tabelle 9 – Vergleich von Stichprobe 2 und 3 in Umfrage I

Stichprobe 2 (N=144) wurde mit Stichprobe 3 (N=195) bezüglich Alter, Geschlechterverteilung, kognitiver Prüfungsangst (G-CTAS I), Sorgen vor dem Physikum (Sorgen I), Lernaufwand (Lernaufwand I), erwarteten Prüfungsleistungen und Abiturnoten verglichen. Dargestellt sind Mittelwerte (M) mit Standardabweichungen (SD). In der Kategorie „Geschlechterverteilung“ sind die Anteile der weiblichen und männlichen Teilnehmer prozentual dargestellt. Abgebildet sind außerdem Beträge der mittleren Differenzen (Δ) von Mittelwerten bzw. der prozentualen Verteilung zwischen Stichprobe 2 und 3. Da es sich bei der Abiturnote um eine freiwillige Angabe handelte, weichen die N-Zahlen vom ursprünglichen Stichprobenumfang ab. Signifikante Unterschiede sind mit * ($p < 0,05$), ** ($p < 0,01$) bzw. *** ($p < 0,001$) gekennzeichnet.

4.4 Stabilität kognitiver Prüfungsangst

Die Überprüfung der Stabilität wurde an der Kontrollgruppe durchgeführt.

Weder in der gesamten Stichprobe noch innerhalb der weiblichen oder männlichen Teilnehmergruppe waren die Unterschiede zwischen G-CTAS I und II signifikant. Mittelwerte sind in Tabelle 10 abgebildet.

G-CTAS I korrelierte stark und sehr signifikant mit G-CTAS II (gesamte Stichprobe (N=92): $r_p=0,75$, $p<0,01$; weibliche Teilnehmer (N=60): $r_p=0,72$, $p<0,01$; männliche Teilnehmer (N=32): $r_p=0,79$ $p<0,01$).

	G-CTAS I (M ± SD)	G-CTAS II (M ± SD)	Δ
Kontrollgruppe (N=92)	22,87 ± 6,96	22,91 ± 6,89	0,04
- weiblich (N=60)	23,68 ± 6,96	23,83 ± 7,04	0,15
- männlich (N=32)	21,34 ± 6,80	21,19 ± 6,35	0,16

Tabelle 10 - Mittelwertunterschiede von G-CTAS I und II

Dargestellt sind Mittelwerte (M) mit Standardabweichung (SD) von G-CTAS-Werten in Umfrage I (G-CTAS I) und II (G-CTAS II) mit Beträgen der mittleren Differenzen(Δ). Werte sind angegeben für die gesamte Kontrollgruppe sowie getrennt für weibliche und männliche Teilnehmer. Signifikante Unterschiede sind mit * ($p<0,05$), ** ($p<0,01$) bzw. *** ($p<0,001$) gekennzeichnet.

4.5 Ausmaß von Sorgen vor dem Physikum und mögliche Zusammenhänge

Die Analysen wurden an Stichprobe 2 durchgeführt. Sorgen vor dem mündlichen Physikum waren im t-Test für verbundene Stichproben hoch signifikant höher als vor dem schriftlichen Physikum ($p < 0,001$; mittlere Differenz: -1,40, T: -16,78, df: 143, siehe Tabelle 11). Die Effektstärke war groß ($d = 1,392$). Histogramme zu Sorgen vor dem schriftlichen und mündlichen Physikum finden sich in Abbildung 9 und 10. 61,8% der Teilnehmer stimmten der Aussage voll zu, Sorgen vor dem mündlichen Physikum zu haben, während der Anteil beim schriftlichen Physikum 9,0 % betrug.

Sorgen vor dem mündlichen Physikum unterschieden sich zwischen weiblichen und männlichen Teilnehmern sehr signifikant ($p < 0,003$, mittlere Differenz: 0,37, T: 3,07, df: 142). Die Effektstärke war moderat ($d = 0,523$). Vor dem schriftlichen Physikum waren die Unterschiede nicht signifikant.

	Sorgen I (schriftlich) (M ± SD)	Sorgen I (mündlich) (M ± SD)
Stichprobe 2 (N=144)	2,10 ± 0,90	3,49 ± 0,73
- weiblich (N=90)	2,20 ± 0,89	3,63 ± 0,66
- männlich (N=54)	1,93 ± 0,89	3,26 ± 0,78

Tabelle 11 – Ausmaß von Sorgen vor schriftlichem und mündlichem Physikum im Vergleich

Stichprobe 2 (N=144). Dargestellt sind Mittelwerte (M) mit Standardabweichung (SD) von Sorgen in Umfrage I vor dem schriftlichen und mündlichen Physikum. Werte sind angegeben für die gesamte Kontrollgruppe sowie getrennt für weibliche und männliche Teilnehmer.

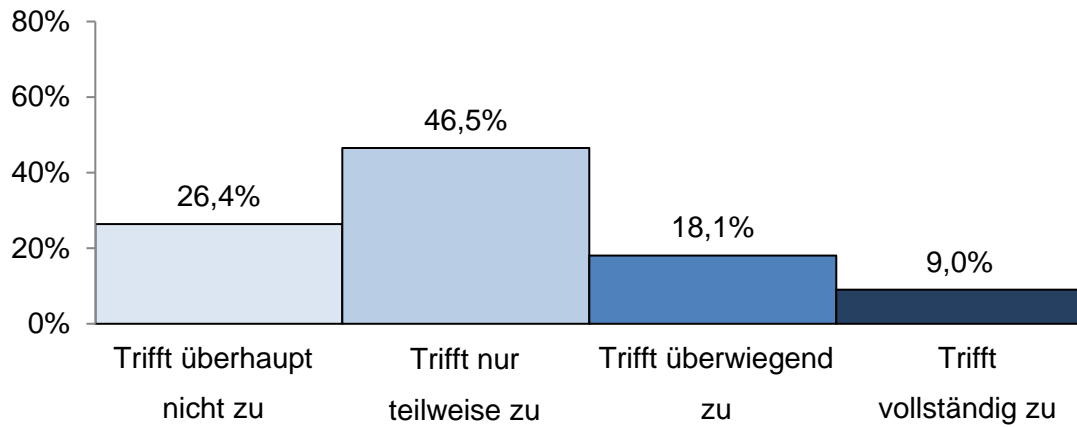


Abbildung 9 – Histogramm „Das schriftliche Physikum bereitet mir Sorgen“

Stichprobe 2 (N=144).

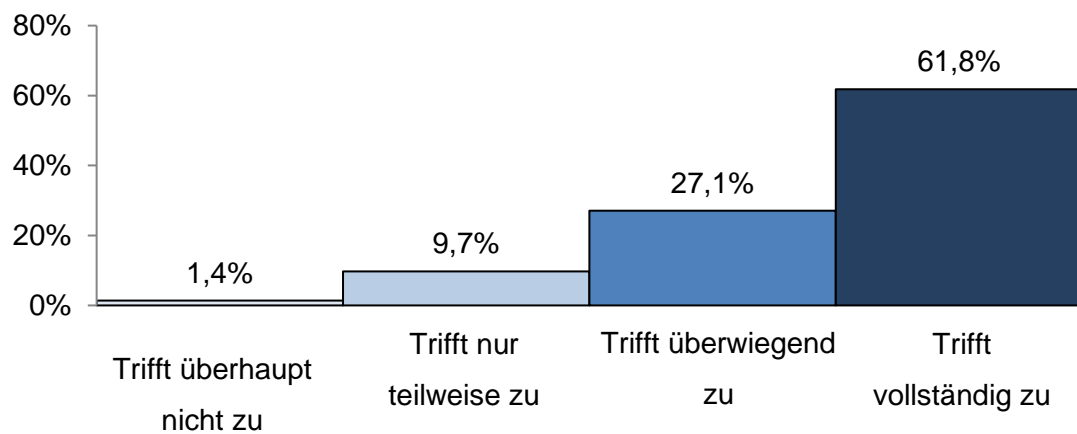


Abbildung 10 – Histogramm „Das mündliche Physikum bereitet mir Sorgen“

Stichprobe 2 (N=144).

Sorgen vor dem schriftlichen (Sorgen I (schriftlich)) und mündlichen Physikum (Sorgen I (mündlich)) in Umfrage I korrelierten moderat und sehr signifikant mit kognitiver Prüfungsangst (G-CTAS I). Sorgen vor dem schriftlichen Physikum korrelierten nicht mit erwarteten Prüfungsleistungen. Sorgen vor dem mündlichen Physikum korrelierten moderat und sehr signifikant mit der erwarteten Prüfungsleistung im mündlichen Physikum, sie korrelierten nicht mit der erwarteten Prüfungsleistung im schriftlichen Physikum. Lernaufwand in Umfrage I (Lernaufwand I), reale Prüfungsleistungen und Abiturnoten korrelierten nicht mit Sorgen vor dem schriftlichen oder mündlichen Physikum.

Unter den weiblichen Teilnehmern korrelierten Sorgen für mündliches und schriftliches Physikum allein mit kognitiver Prüfungsangst (G-CTAS I) signifikant. Unter den männlichen Teilnehmern zeigten sich keine relevanten Korrelationen.

Alle Werte sind in Tabelle 12 abgebildet.

	Sorgen I (schriftlich)	Sorgen I (mündlich)
gesamte Stichprobe (N=144)		
G-CTAS I	0,35**	0,33**
Lernaufwand I	0,02	0,11
Erwarteten Prüfungsleistungen		
- schriftliches Physikum	0,28**	0,05
- mündliches Physikum	0,09	0,31**
Reale Prüfungsleistungen (N=135)		
- schriftliches Physikum	0,16	0,06
- mündliches Physikum	0,14	0,00
Abiturnote (N=140)	0,20	-0,07
Weibliche Teilnehmer (N=90)		
G-CTAS I	0,39**	0,39**
Lernaufwand I	-0,09	-0,06
Erwarteten Prüfungsleistungen		
- schriftliches Physikum	0,29**	0,03
- mündliches Physikum	0,04	0,27**
Reale Prüfungsleistungen (N=85)		
- schriftliches Physikum	0,16	0,05
- mündliches Physikum	0,13	-0,01
Abiturnote (N=88)	0,12	0,04

Männliche Teilnehmer (N=54)		
G-CTAS I	0,25	0,19
Lernaufwand I	0,12	0,24
Erwarteten Prüfungsleistungen		
- schriftliches Physikum	0,23	0,06
- mündliches Physikum	0,10	0,28*
Reale Prüfungsleistungen (N=50)		
- schriftliches Physikum	0,15	0,06
- mündliches Physikum	0,18	0,04
Abiturnote (N=52)	0,24	-0,29*

Tabelle 12 – Korrelationen mit Sorgen vor schriftlichem und mündlichem Physikum

Stichprobe 2, N=144. Sorgen vor dem schriftlichen (Sorgen I (schriftlich)) und mündlichen Physikum (Sorgen (mündlich)) wurden mit kognitiver Prüfungsangst, Lernaufwand, erwarteten Prüfungsleistungen, realen Prüfungsleistungen und bisherigen Prüfungsleistungen (Abiturnote) korreliert (r_p). Die Berechnungen wurden für die gesamte Stichprobe sowie getrennt für weibliche und männliche Teilnehmer durchgeführt. Da es sich bei realen Prüfungsleistungen und Abiturnoten um freiwillige Angaben handelte, weichen die N-Zahlen von den ursprünglichen Stichprobenumfängen ab. Signifikante Unterschiede sind mit * ($p < 0,05$), ** ($p < 0,01$) bzw. *** ($p < 0,001$) gekennzeichnet.

4.6 Examenssimulationen

Die folgenden Analysen wurden durch Vergleiche von Simulations- (N=52) und Kontrollgruppe (N=92) durchgeführt.

4.6.1 Zielgruppe der Examenssimulation

Alter, Geschlechterverteilung, kognitive Prüfungsangst (G-CTAS I), Sorgen vor dem schriftlichen und mündlichen Physikum (Sorgen I), Lernaufwand (Lernaufwand I), erwartete Prüfungsleistungen im schriftlichen und mündlichen Physikum sowie Abiturnoten unterschieden sich nicht signifikant zwischen Simulations- und Kontrollgruppe (*siehe Tabelle 13*).

	Simulationsgruppe (M ± SD)	Kontrollgruppe (M ± SD)	Δ
Alter	21,90 ± 3,29	21,99 ± 3,09	0,09
G-CTAS I	22,46 ± 6,21	22,87 ± 6,96	0,41
Sorgen I			
- schriftliches Physikum	2,19 ± 0,93	2,04 ± 0,88	0,15
- mündliches Physikum	3,35 ± 0,81	3,58 ± 0,67	0,23
Lernaufwand I	3,21 ± 0,70	3,24 ± 0,76	0,03
Erwartete Prüfungsleistungen			
- schriftliches Physikum	2,46 ± 0,67	2,41 ± 0,60	0,05
- mündliches Physikum	2,73 ± 0,77	2,92 ± 0,84	0,19
Abiturnote (Simulationsgruppe: N=51, Kontrollgruppe: N=89)	1,51 ± 0,49	1,50 ± 0,52	0,01
Geschlechterverteilung (prozentual)			
- weiblich	57,7%	65,2%	7,5%
- männlich	42,3%	34,8%	7,5%

Tabelle 13 - Vergleich von Simulations- und Kontrollgruppe in Umfrage I

Simulationsgruppe, N=52; Kontrollgruppe, N=92. Dargestellt sind Mittelwerte (M) mit Standardabweichung (SD) von Alter, G-CTAS-Werten (G-CTAS I), Sorgen vor dem Physikum (Sorgen I), Lernaufwand (Lernaufwand I), erwarteten Prüfungsleistungen und bisherigen Prüfungsleistungen (Abiturnote) sowie die prozentuale Verteilung der Geschlechter getrennt für beide Gruppen. Dargestellt sind außerdem Beträge der mittleren Differenzen (Δ). Da es sich bei der Abiturnote um keine verpflichtende Angabe handelte, weichen die N-Zahlen von den ursprünglichen Stichprobenumfängen ab. Signifikante Unterschiede sind mit * (p<0,05), ** (p<0,01) bzw. *** (p<0,001) gekennzeichnet.

4.6.2 Effekte der Examenssimulation

In Tabelle 14 sind Mittelwerte mit Mittelwertunterschieden der einzelnen Untersuchungskategorien für Simulations- und Kontrollgruppe dargestellt. Abgebildet sind die Werte außerdem getrennt für weibliche und männliche Teilnehmer. Männliche Teilnehmer der Simulationsgruppe schnitten signifikant schlechter im mündlichen Physikum ab als Teilnehmer der Kontrollgruppe ($U=211,500$, $p=0,049$). Die Effektstärke war klein ($d=-0,496$). In den übrigen Analysen konnten keine signifikanten Unterschiede zwischen Simulations- und Kontrollgruppe festgestellt werden.

	Simulationsgruppe (M ± SD)	Kontrollgruppe (M ± SD)	Δ
gesamte Stichprobe (Simulationsgruppe: N=52, Kontrollgruppe: N=92)			
G-CTAS II	22,12 ± 5,91	22,91 ± 6,89	0,81
Sorgen II			
- schriftliches Physikum	2,35 ± 0,93	2,17 ± 1,10	0,17
- mündliches Physikum	3,50 ± 0,70	3,43 ± 1,06	0,07
Lernaufwand II (Simulationsgruppe: N=52, Kontrollgruppe: N=84)			
- schriftliches Physikum	3,21 ± 0,72	3,13 ± 0,76	0,08
- mündliches Physikum	3,29 ± 0,72	3,35 ± 0,78	0,06
Reale Prüfungsleistungen (Simulationsgruppe: N=52, Kontrollgruppe: N=83)			
- schriftliches Physikum	2,52 ± 0,92	2,52 ± 0,83	0,00
- mündliches Physikum	2,67 ± 1,02	2,48 ± 1,05	0,19
weibliche Teilnehmer (Simulationsgruppe: N=30, Kontrollgruppe: N=60)			
G-CTAS II	22,43 ± 5,02	23,83 ± 7,04	1,40
Sorgen II			
- schriftliches Physikum	2,37 ± 1,03	2,28 ± 1,06	0,08
- mündliches Physikum	3,63 ± 0,67	3,57 ± 0,98	0,07
Lernaufwand II (Simulationsgruppe: N=30, Kontrollgruppe: N=55)			
- schriftliches Physikum	3,20 ± 0,71	3,22 ± 0,63	0,02
- mündliches Physikum	3,43 ± 0,68	3,44 ± 0,74	0,01
Reale Prüfungsleistungen (Simulationsgruppe: N=30, Kontrollgruppe: N=55)			
- schriftliches Physikum	2,40 ± 0,89	2,62 ± 0,78	0,22
- mündliches Physikum	2,50 ± 1,14	2,53 ± 0,98	0,03

männliche Teilnehmer (Simulationsgruppe: N=22, Kontrollgruppe: N=32)			
G-CTAS II	21,68 ± 7,05	21,19 ± 6,35	0,49
Sorgen II			
- schriftliches Physikum	2,32 ± 0,78	1,97 ± 1,15	0,35
- mündliches Physikum	3,32 ± 0,72	3,19 ± 1,18	0,13
Lernaufwand II (Simulationsgruppe: N=22, Kontrollgruppe: N=29)			
- schriftliches Physikum	3,23 ± 0,75	2,97 ± 0,94	0,26
- mündliches Physikum	3,09 ± 0,75	3,17 ± 0,85	0,08
Reale Prüfungsleistungen (Simulationsgruppe: N=22, Kontrollgruppe: N=28)			
- schriftliches Physikum	2,68 ± 0,95	2,32 ± 0,91	0,36
- mündliches Physikum	2,91 ± 0,81	2,39 ± 1,20	0,52*

Tabelle 14 - Vergleich von Simulations- und Kontrollgruppe in Umfrage II

Simulationsgruppe, N=52; Kontrollgruppe, N=92. Dargestellt sind Mittelwerte (M) mit Standardabweichung (SD) von G-CTAS-Werten in Umfrage II (G-CTAS II), retrospektiven Sorgen vor dem schriftlichen und mündlichen Physikum (Sorgen II), retrospektivem Lernaufwand auf das schriftliche und mündliche Physikum (Lernaufwand II) und realen Prüfungsleistungen im schriftlichen und mündlichen Physikum. Dargestellt sind zudem Beträge der mittleren Differenzen (Δ). Alle Werte sind für die gesamte Stichprobe sowie getrennt für weibliche und männliche Teilnehmer abgebildet. Da es sich bei retrospektivem Lernaufwand und realen Prüfungsleistungen um keine verpflichtenden Angaben handelte, weichen die N-Zahlen von den ursprünglichen Stichprobenumfängen ab. Signifikante Unterschiede sind mit * ($p < 0,05$), ** ($p < 0,01$) bzw. *** ($p < 0,001$) gekennzeichnet.

Mittelwerte mit Standardabweichungen von G-CTAS I und II sowie deren Unterschiede sind für Simulations- und Kontrollgruppe in Tabelle 15 angegeben. Dargestellt sind die Werte für die gesamten Stichproben sowie getrennt nach Geschlechtern. Weder in der Simulations- noch in der Kontrollgruppe konnten signifikante Unterschiede der G-CTAS-Werte im t-Test bzw. im Wilcoxon-Test für verbundene Stichproben festgestellt werden.

	G-CTAS I	G-CTAS II	
	M ± SD	M ± SD	Δ
Simulationsgruppe (N=52)	22,46 ± 6,21	22,12 ± 5,91	0,34
- weiblich (N=30)	23,10 ± 6,26	22,43 ± 5,02	0,67
- männlich (N=22)	21,59 ± 6,19	21,68 ± 7,05	0,09
Kontrollgruppe (N=92)	22,87 ± 6,96	22,91 ± 6,89	0,04
- weiblich (N=60)	23,68 ± 6,96	23,83 ± 7,04	0,15
- männlich (N=32)	21,34 ± 6,80	21,19 ± 6,35	0,15

Tabelle 15 – Unterschiede von G-CTAS-Werten in Umfrage I und II

*Angegeben sind Mittelwerte (M) mit Standardabweichung (SD) von Werten in Umfrage I und II sowie Beträge der mittleren Differenzen (Δ). Alle Werte sind für die gesamte Stichprobe sowie getrennt für weibliche und männliche Teilnehmer abgebildet. Signifikante Unterschiede sind mit * ($p < 0,05$), ** ($p < 0,01$) bzw. *** ($p < 0,001$) gekennzeichnet.*

Ebenfalls konnten im t-Test bzw. im Wilcoxon-Test für verbundene Stichproben in keiner Gruppe signifikante Unterschiede zwischen Sorgen I und Sorgen II vor schriftlichem oder mündlichem Physikum festgestellt werden. Mittelwerte mit Standardabweichungen und mittlere Differenzen sind für Simulations- und Kontrollgruppe sowie getrennt nach Geschlechtern in Tabelle 16 für schriftliches und mündliches Physikum dargestellt.

	Sorgen I	Sorgen II	Δ
	Schriftliches Physikum		
Simulationsgruppe (N=52)	2,19 \pm 0,93	2,35 \pm 0,93	0,16
- weiblich (N=30)	2,27 \pm 0,91	2,37 \pm 1,03	0,10
- männlich (N=22)	2,09 \pm 0,97	2,32 \pm 0,78	0,23
Kontrollgruppe (N=92)	2,04 \pm 0,88	2,17 \pm 1,10	0,13
- weiblich (N=60)	2,17 \pm 0,89	2,28 \pm 1,06	0,11
- männlich (N=32)	1,81 \pm 0,82	1,97 \pm 1,15	0,16
	Mündliches Physikum		
Simulationsgruppe (N=52)	3,35 \pm 0,81	3,50 \pm 0,70	0,15
- weiblich (N=30)	3,50 \pm 0,78	3,63 \pm 0,67	0,13
- männlich (N=22)	3,14 \pm 0,83	3,32 \pm 0,72	0,18
Kontrollgruppe (N=92)	3,58 \pm 0,67	3,43 \pm 1,06	0,15
- weiblich (N=60)	3,70 \pm 0,59	3,57 \pm 0,98	0,13
- männlich (N=32)	3,34 \pm 0,75	3,19 \pm 1,18	0,15

Tabelle 16 - Unterschiede von Sorgen vor schriftlichem und mündlichem Physikum in Umfrage I und II

Angegeben sind Mittelwerte (M) mit Standardabweichung (SD) von Werten in Umfrage I und II. Zudem sind Beträge der mittleren Differenzen (Δ) dargestellt. Alle Werte sind für die gesamte Stichprobe sowie getrennt für weibliche und männliche Teilnehmer abgebildet. Signifikante Unterschiede sind mit * ($p < 0,05$), ** ($p < 0,01$) bzw. *** ($p < 0,001$) gekennzeichnet.

Erwartete Prüfungsleistungen im schriftlichen Physikum unterschieden sich im t-Test bzw. im Wilcoxon-Test für verbundene Stichproben weder in der Simulations- noch in der Kontrollgruppe von realen Prüfungsleistungen im schriftlichen Physikum. Dies war auch in den Untersuchungen getrennt nach Geschlechtern der Fall.

Erwartete Prüfungsleistungen unterschieden sich teilweise von realen Prüfungsleistungen im mündlichen Physikum: Männliche Teilnehmer der Simulationsgruppe schnitten signifikant schlechter ab als erwartet ($z=-2,14$, $p=0,033$, $N=22$), die Effektstärke war klein ($d=0,497$). Reale Prüfungsleistungen im mündlichen Physikum waren in der Kontrollgruppe hoch signifikant besser als erwartete Prüfungsleistungen ($p<0,001$, mittlere Differenz: $-0,48$; $T:4,00$; $df: 82$), die Effektstärke war klein ($d=-0,441$). Unterschiede der weiblichen Teilnehmer der Kontrollgruppe waren dabei hoch signifikant ($p<0,001$; mittlere Differenz: $0,55$; $T:4,14$; $df: 54$) bei moderater Effektstärke ($d=-0,556$). Unterschiede der männlichen Teilnehmer waren nicht signifikant. Alle Werte sind in Tabelle 17 abgebildet.

	Erwartete Prüfungsleistung	Reale Prüfungsleistung	Δ
Schriftliches Physikum			
Simulationsgruppe (N=52)	2,46 \pm 0,67	2,52 \pm 0,92	0,06
- weiblich (N=30)	2,47 \pm 0,68	2,40 \pm 0,89	0,07
- männlich (N=22)	2,45 \pm 0,67	2,68 \pm 0,95	0,23
Kontrollgruppe (N=83)	2,42 \pm 0,59	2,52 \pm 0,83	0,10
- weiblich (N=55)	2,44 \pm 0,63	2,62 \pm 0,78	0,18
- männlich (N=28)	2,39 \pm 0,50	2,32 \pm 0,91	0,07
Mündliches Physikum			
Simulationsgruppe (N=52)	2,73 \pm 0,77	2,67 \pm 1,02	0,06
- weiblich (N=30)	2,87 \pm 0,78	2,50 \pm 1,14	0,37
- männlich (N=22)	2,55 \pm 0,74	2,91 \pm 0,81	0,36*
Kontrollgruppe (N=83)	2,96 \pm 0,86	2,48 \pm 1,05	0,48***
- weiblich (N=55)	3,07 \pm 0,81	2,53 \pm 0,98	0,54***
- männlich (N=28)	2,75 \pm 0,93	2,39 \pm 1,20	0,36

Tabelle 17 – Unterschiede von erwarteten Prüfungsleistungen und realen Prüfungsleistungen im schriftlichen und mündlichen Physikum

Angegeben sind Mittelwerte (M) mit Standardabweichung (SD) von Werten in Umfrage I und II. Zudem sind Beträge der mittleren Differenzen (Δ) dargestellt. Alle Werte sind für die gesamte Stichprobe sowie getrennt für weibliche und männliche Teilnehmer abgebildet. Signifikante Unterschiede sind mit * ($p < 0,05$), ** ($p < 0,01$) bzw. *** ($p < 0,001$) gekennzeichnet.

5 Diskussion

5.1 Diskussion des Materials und der Methoden

5.1.1 Zusammenfassung des Materials und der Methoden

Die Datenerhebung erfolgte mittels online-Fragebogenstudie von Herbst 2015 bis Herbst 2016 im Vorher-nachher-Vergleich jeweils wenige Wochen vor bzw. nach dem schriftlichen Physikum unter Studierenden der Humanmedizin an der LMU München. Nach Versenden der ersten Umfrage wurde allen interessierten Physikumskandidaten die Teilnahme an einer freiwilligen Examenssimulation ermöglicht. In Abhängigkeit von der Bearbeitung der Umfragen und Teilnahme an der Examenssimulation erfolgte die Zuteilung zu den Stichproben.

In der vorliegenden Arbeit wurde erstmals eine deutsche Version der „Cognitive Test Anxiety Scale“ (CTAS) erstellt, mittels Item- und Testanalyse sowie exploratorischer Faktorenanalyse (EFA) überprüft und validiert. Mit dem auf diese Weise erstellten Messinstrument (G-CTAS) wurde das Ausmaß kognitiver Prüfungsangst unter Medizinstudierenden unmittelbar vor dem Physikum untersucht. Dabei wurde außerdem auf Unterschiede zwischen den Geschlechtern eingegangen. Um einen Selektionsbias zwischen Teilnehmern beider Umfragen mit Teilnehmern lediglich der ersten Umfrage in den Vorher-nachher-Untersuchungen ausschließen zu können, wurden diese in den zu untersuchenden Punkten miteinander verglichen. Die kognitive Prüfungsangst wurde durch Vergleich von G-CTAS-Werten vor und nach dem Physikum innerhalb der Kontrollgruppe auf ihre Stabilität überprüft. Zum besseren Verständnis der durch das Physikum entstehenden Stressbelastung wurden schriftliches und mündliches Physikum bezüglich des Ausmaßes an verursachenden Prüfungssorgen miteinander verglichen und mögliche Zusammenhänge mit kognitiver Prüfungsangst, Lernaufwand, erwarteten Prüfungsleistungen, realen Prüfungsleistungen und Abiturnoten mittels Korrelationen auf ihre Relevanz analysiert. Die Untersuchung wurde außerdem getrennt für weibliche und männliche Teilnehmer durchgeführt.

Darüber hinaus wurden Examenssimulationen auf ihre Eignung als sinnvolles Prüfungsvorbereitungsangebot überprüft. Dies geschah durch Analyse der Zielgruppe der Examenssimulation sowie Untersuchung von Effekten auf kognitive Prüfungsangst, Sorgen vor dem Physikum, Lernaufwand und reale Prüfungsleistungen.

5.1.2 Aufbau der Studie

5.1.2.1 Studiendesign

5.1.2.1.1 Durchführung als online-Studie

In der vorliegenden Arbeit wurde wie von Philips, Reddy et al. [115] empfohlen die Versendung eines Anschreibens per E-Mail mit Link zu einer Umfragenplattform gewählt. Im Vergleich zum Papierformat bringt eine Online-Befragung bezüglich Kosten, flexibler Durchführbarkeit, Anonymität und entfallender Dateneingabe erhebliche Vorteile mit sich und hat sich aus diesen Gründen in den vergangenen Jahren zum populären Datenerhebungsinstrument entwickelt [116]. Im Gegensatz zur Durchführung auf Papier-Format führen online-Befragungen bei Studierenden jedoch zu geringeren Rücklaufquoten [117, 118]. Alternativen hätten eine verpflichtende Teilnahme oder die Durchführung auf Papier-Format dargestellt, beispielsweise direkt vor der realen schriftlichen Physikumsprüfung oder im Rahmen eines Vorbereitungsseminars. Da zum geplanten Zeitpunkt der Studiendurchführung keine Seminare von der Universität angeboten wurden, ergab sich diese Möglichkeit nicht. Eine Befragung unmittelbar vor der realen schriftlichen Prüfung kam nicht in Frage, da ein Teil der Studierenden das mündliche Physikum bereits absolviert hatte und dies die Auswertung maßgeblich erschwert hätte. Gegen eine verpflichtende Umfragen-Teilnahme entschlossen wir uns, da dies im Vorfeld einer stark belastenden Situation wie dem Physikum nicht angebracht erschien.

5.1.2.1.2 Durchführungs- und Befragungszeitpunkt

Mehrere Punkte sprachen für die Durchführung von Befragung und Examenssimulation möglichst nah am Termin des realen schriftlichen Physikums, beispielsweise das Ziel der Messung der kognitiven Prüfungsangst unmittelbar vor der realen Prüfung, ein möglichst vollständiger Lernstand der Teilnehmer sowie die wahrscheinlichere Präsenz von Prüfungssorgen.

Dementgegen standen Überlegungen, welche eine Durchführung weiter im Vorfeld der realen Prüfung sinnvoll erscheinen ließen. So wurde die Rücklaufquote als potentiell höher eingeschätzt als unmittelbar vor der Prüfung. Auch das Antwortverhalten sollte nicht durch die bereits erfolgte Teilnahme am mündlichen Physikum verfälscht werden. Nach Teilnahme an der Examenssimulation sollte zudem ausreichend Zeit bleiben, auf dort identifizierte Lücken in der Vorbereitung einzugehen.

Mit den gewählten Durchführungszeitpunkten wurde in Bezug auf die erwähnten Punkte ein Kompromiss eingegangen, durch welchen sich allerdings ein möglicher Einfluss des mündlichen Prüfungstermins auf die Ergebnisse und Teilnehmer der Examenssimulationen ergab. Zu allen drei Durchführungszeitpunkten wurden die Termine für das mündliche Physikum während der Befragung mittels Umfrage I

veröffentlicht. Es ist somit möglich, dass der mündliche Prüfungstermin einen Einfluss unter anderem auf die kognitive Prüfungsangst und Prüfungssorgen der Teilnehmer gehabt haben könnte. Zum Zeitpunkt der Examenssimulation hatten die mündlichen Prüfungen bereits begonnen, sodass einige Studierende aufgrund einer Terminüberschneidung nicht an den Examenssimulationen teilnehmen konnten.

5.1.2.1.3 Sponsoring der Examenssimulationen durch einen Verlag

Prüfungsunterlagen der Examenssimulationen wurden den Teilnehmern durch den Verlag „Elsevier“ kostenlos zur Verfügung gestellt. Dadurch konnte der Verlag mittels Flyern und Gestaltung der Hefte für sich werben, die Zusammenarbeit war jedoch nicht mit der Weitergabe von Teilnehmerdaten verbunden. Durch die kostenlose Bereitstellung der Hefte konnte allen interessierten Studierenden die Teilnahme an den Examenssimulationen ermöglicht werden. Dies wäre ohne das Sponsoring aus Kostengründen nicht möglich gewesen.

Die Kooperation mit dem Verlag brachte außerdem einen weiteren Vorteil mit sich: In Rücksprache wurde die Idee des Angebots von Prüfungssimulationen durch den Verlag verbreitet, woraufhin Simulationen des schriftlichen Physikums von Sommer 2015 bis Sommer 2016 an weiteren medizinischen Fakultäten in Deutschland durchgeführt wurden (darunter 10 Universitäten mit insgesamt über 1300 Teilnehmern im Sommer 2015).

5.1.2.2 Datenschutz

Wie beschrieben war die Teilnahme an Umfragen und Examenssimulationen freiwillig. Falls Teilnehmer nicht mit der Angabe ihrer persönlichen Daten, insbesondere der Matrikelnummer, in den Umfragen oder der Examenssimulation einverstanden waren, konnten sie die Teilnahme abbrechen, ohne dass dies mit Nachteilen verbunden war.

Die Generierung eines Verschlüsselungscodes auf Basis der Matrikelnummer erlaubte eine einfache und eindeutige Zuordnung für den Vorher-nachher-Vergleich sowie zu Simulations- und Kontrollgruppe. Es handelte sich dabei um ein pseudonymisiertes Verfahren. Der Zugang zum Verschlüsselungscode war auf die Studienleitung, stellvertretende Studienleitung und Doktorandin beschränkt. Dieses Vorgehen war von der Ethikkommission genehmigt worden (*siehe 3.1.3 Ethikvotum*).

Inzwischen wurde ein alternatives, ebenfalls pseudonymisiertes Verfahren etabliert: Mithilfe mehrerer eindeutig zu beantwortender Fragen (Beispiel: Erster Buchstabe des Vornamens Ihres Vaters) kann ein individueller Code generiert und für die Zuordnung verwendet werden. Eine Identifikation der Teilnehmer ist hierbei weder für Studienleitung, stellvertretende Studienleitung oder Doktoranden möglich.

5.1.3 Messinstrumente

5.1.3.1 Verwendung von CTAS

Obwohl im Deutschen bereits mehrere Fragebögen zur Messung von Prüfungsangst etabliert sind, wurde für diese Arbeit ein englischer Fragebogen neu übersetzt. Es wurde außerdem die ursprüngliche Version des Fragebogens (CTAS) [67] verwendet.

Aufgrund der weitreichenden negativen Folgen stand die Messung der kognitiven Komponente im Vordergrund der Untersuchungen. Mögliche Alternativen hätten die Verwendung des TAI oder RTT dargestellt, da sie ebenfalls Items zur Messung der kognitiven Komponente enthalten. Beide Fragebögen sind zudem in deutscher Fassung verfügbar (*siehe 1.1.7 Messinstrumente von Prüfungsangst*).

Mithilfe des TAI-G ist ein genaues Verständnis der Auswirkungen kognitiver Prüfungsangst in der Vorbereitungs- und Reflexionsphase einer Prüfung allerdings nicht möglich, da lediglich die Prüfungsangst während der Prüfungssituation selbst gemessen wird. Auch sind die für das TAI vorliegenden Normwerte laut Szafranski, Barrera et al. [119] bereits überholt. Erwähnenswert ist zudem, dass kein kostenfreier Zugang zu den verschiedenen Versionen des TAI-G verfügbar ist.

Die deutsche Übersetzung des RTT [65] enthält insgesamt 33 Items, worunter lediglich sieben Items zur Skala „worry“ und neun Items zur Skala „test irrelevant thinking“ gezählt werden (zusammengefasst von Rost & Schermer [120]). Die Identifikation der Subskalen erbrachte außerdem „nur unklare Befunde“ [120]. RTT ist aus diesen Gründen ebenfalls nur bedingt geeignet, speziell kognitive Prüfungsangst zu messen und zu verstehen.

Stattdessen wurde für die Verwendung von CTAS und deren Übersetzung ins Deutsche entschieden, da dieser Fragebogen speziell zur Messung der kognitiven Prüfungsangstkomponente entwickelt wurde [67]. Im Januar 2015 wurde die Skala ins Deutsche übersetzt, zu diesem Zeitpunkt war bereits bekannt, dass die invers gepolten Items einen zweiten Faktor („test confidence“) erzeugen [75]. Cassady & Finch hatten deshalb eine Überprüfung des Fragebogens vorgenommen und drei alternative Modelle verglichen [77]. Obwohl sie die Gültigkeit der ursprünglichen Skala garantierten, empfahlen sie aufgrund überlegener Statistikkennwerte die Verwendung einer eindimensionalen Kurzform, welche aus 17 nicht invers gepolten Items besteht.

Für diese Arbeit wurde die ursprüngliche 27-Items Skala (CTAS) übersetzt und verwendet, da sich daraus die Kurzform einfach generieren lässt. Auch bringt die Verwendung der in dieser Version enthaltenen invers gepolten Items Vorteile mit sich. Beispielsweise zwingen sie den Leser, Items genauer zu lesen [121] und können eine Verfälschung durch die soziale Erwünschtheit von Antworten reduzieren [122].

Im Juni 2015 stellten Cassady & Finch eine Überarbeitung der Skala vor (CTAR) [76]. Diese besteht aus 25 Items, wovon wiederum die oben genannten 17 Items dem Original-Fragebogen entstammen, sowie acht weiteren Items. Mittlerweile existiert auch

eine Kurzform der Skala [79]. Während die 17 Items des Originalfragebogens regelmäßig Bestandteil der Skala sind (bzw. zwölf Items in der Kurzform), unterliegen die für CTAR neu entwickelten acht Items bis zum aktuellen Zeitpunkt einem Überarbeitungsprozess und wurden zuletzt 2018 [78] teilweise wieder entfernt.

Die Verwendung der ursprünglichen Version ist deshalb auch im Rückblick nicht als Einschränkung dieser Arbeit zu sehen.

5.1.3.2 Übersetzung ins Deutsche

Bei der Übersetzung der englischen Items ins Deutsche wurde der Fokus primär auf eine inhaltsgenaue, weniger auf eine semantisch korrekte Übersetzung gelegt. Die Methodik des angewandten Vorwärts-Rückwärts-Verfahrens war etabliert und bereits von mehreren Forschergruppen erfolgreich angewendet worden [75, 79, 80, 123, 124].

Im Übersetzungsprozess war Item 13 („Mir fallen Tests unter Zeitdruck besonders leicht.“) wiederholt diskutiert worden, da die deutsche Übersetzung der Bedeutung der englischen Original-Frage nicht vollständig entsprach. Dass Item 13 in der Item- und Testanalyse ungenügende Kennwerte zeigte, könnte auf die nicht optimal gelungene Übersetzung zurückzuführen sein. Nur sehr wenige Teilnehmer stimmten der invers gepolten Aussage des Items zu, die Schwierigkeit und die Varianz des umgepolten Items waren sehr niedrig. Durch Umformulierung des Items hätte man dem möglicherweise entgegenwirken können. Sinnvoll in Hinblick auf eine Erhöhung der Schwierigkeit erscheint beispielsweise die Umformulierung des Items in „Mir fallen Tests unter Zeitdruck nicht schwer“. Diese Formulierung war jedoch aufgrund einer zu starken Abweichung von der englischen Original-Frage verworfen worden.

5.1.3.3 Zusatzmaterial

Der Fokus dieser Arbeit lag auf der Messung und Analyse kognitiver Prüfungsangst. Die Zusatzfragen sollten lediglich einen Überblick, keine differenzierte Analyse von Zusammenhängen und Auswirkungen ermöglichen. Auf die Verwendung von aufwändigen Messmethoden oder Skalen wurde demzufolge verzichtet, in mehreren Kategorien wurden lediglich Einzelfragen verwendet. Für die Bestimmung von Sorgen und Lernaufwand wurde außerdem auf subjektive Fragen, nicht auf objektive Messmethoden zurückgegriffen. Dies schränkt die Aussagekraft der Analysen bezüglich des Zusatzmaterials ein.

Im Gegensatz zur Messung des Lernaufwands in Umfrage I wurde der retrospektive Lernaufwand in Umfrage II getrennt für schriftliches und mündliches Physikum erhoben. Dies erschien in Umfrage I nicht sinnvoll, da sich Prüfungstermine des mündlichen Physikums unter den Teilnehmern teilweise stark unterschieden und davon ausgegangen wurde, dass der Lernaufwand auf die einzelnen Prüfungen in deren

Abhängigkeit variieren würde (*siehe 1.2.1 Der Erste Abschnitt der Ärztlichen Prüfung (Physikum)*).

5.1.3.4 Antwortformate

Für die Antwortkategorien „Sorgen vor dem Physikum“, „Sorgen vor dem Physikum - retrospektiv“, „Lernaufwand“ und „Lernaufwand - retrospektiv“ wurde in Anlehnung an das Antwortformat von G-CTAS eine vierstufige Likert-Skala verwendet. Vorteilhaft an einer geraden Anzahl von Skalenstufen ist, dass keine neutrale Aussage getroffen werden kann, Teilnehmer mussten sich zumindest für eine Tendenz entscheiden.

Obwohl laut Moosbrugger [125] ab sieben Skalenstufen kein zusätzlicher Informationsgewinn zu erwarten ist, hätte die Ausweitung auf eine sechsstufige Likertskala für eine differenziertere Messung eine sinnvolle Alternative dargestellt. Um einen Überblick über Zusammenhänge und Auswirkungen zu gewinnen, war das gewählte Antwortformat dennoch geeignet.

Ähnlich war dies in den Kategorien „erwartete Prüfungsleistungen“ und „reale Prüfungsleistungen“. Hier wurde mit Schulnoten von 1 (=sehr gut) bis 5 (=mangelhaft) gearbeitet. Da für die mündliche Prüfung keine stärker differenzierte Bewertung vorgenommen wird (*siehe 1.2.1 Der Erste Abschnitt der Ärztlichen Prüfung (Physikum)*), bot sich hier im Antwortformat keine Alternative. In der schriftlichen Prüfung wird den Prüflingen dagegen in der offiziellen Ergebnisbekanntgabe die Prozentzahl der richtig beantworteten Prüfungsfragen mitgeteilt. Da die meisten Studierenden diese Prozentzahl im Gegensatz zur Note im schriftlichen Physikum jedoch nicht lange präsent haben, wurde die zuverlässigere Variante der Angabe in Schulnoten gewählt.

5.1.4 Stichproben

5.1.4.1 Rücklaufquoten und Stichprobenumfänge

Zwischen Herbst 2015 und Herbst 2016 nahmen laut IMPP [126] insgesamt 1589 Studierende der LMU München am schriftlichen Physikum teil. Der Stichprobenumfang von Stichprobe 1 betrug 339 und entsprach somit 21,3% der Zielgruppe.

Laut Sheehan [127] lagen Rücklaufquoten von Umfragen mit vergleichbarem Studiendesign zwischen 1986 und 2000 im Durchschnitt bei 36,8%, gingen jedoch innerhalb dieses Zeitraums zurück. Ähnliche Beobachtungen wurden auch, vor allem unter Studierenden, in anderen Studien beschrieben [128, 129]. Ein möglicher Grund ist in einem Überangebot zu sehen, welches durch häufiges Versenden von Umfragen insbesondere an Universitäten zustande kommt [130].

Eine geschätzte Rücklaufquote von 21,3% kann in Anbetracht der Analyse von Sheehan dennoch als leicht unterdurchschnittlich betrachtet werden. Eine Erklärung könnte sein, dass die Untersuchungen während der Vorbereitungsphase im Vorfeld relevanter Prüfungen durchgeführt wurden.

339 Studierende beantworteten Umfrage I, davon nahmen 42,5% auch an Umfrage II teil (Stichprobe 2, N=144). Die Teilnehmeranzahl der Umfrage II betrug N=284 und fiel somit etwas geringer aus als bei Umfrage I. Dies könnte unter anderem dadurch zu erklären sein, dass Studierende nach dem Physikum die Universität wechselten und nicht mehr über den verwendeten E-Mail-Verteiler zu erreichen waren. Eine weitere mögliche Erklärung könnte sein, dass sich einige Studierende letztendlich doch gegen eine Teilnahme am realen Physikum entschieden und deshalb auf die zweite Umfrageteilnahme verzichteten.

Bei der Betrachtung der Teilnehmerzahlen (*siehe Abbildung 11*) wird deutlich, dass für den Größenunterschied zwischen Stichprobe 1 und 2 nicht primär die geringere Teilnehmerzahl von Umfrage II verantwortlich war, sondern vielmehr das Studiendesign mit Durchführung eines Vorher-nachher-Vergleich.

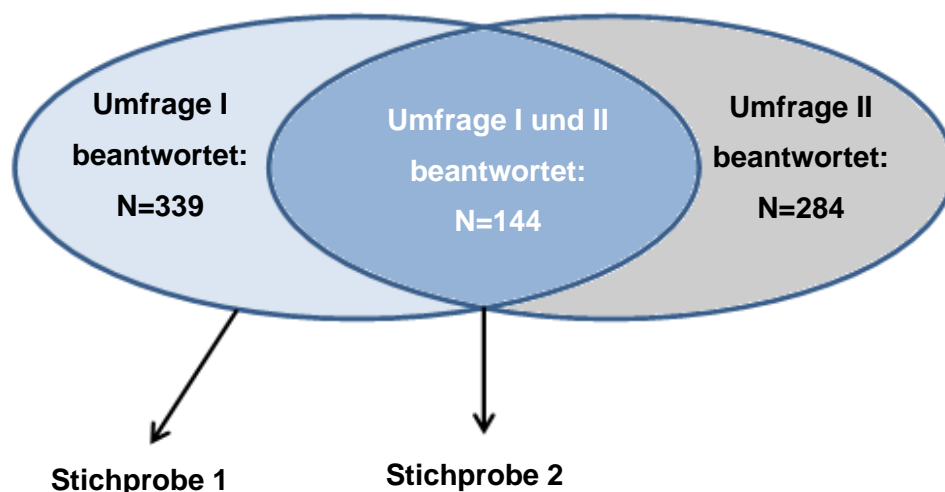


Abbildung 11 – Teilnehmerzahlen Umfrage I und II

339 Teilnehmer hatten Umfrage I vollständig beantwortet (=Stichprobe 1), 284 Teilnehmer Umfrage II. 144 Teilnehmer hatten beide Umfragen vollständig beantwortet (Stichprobe 2).

Von den 144 Teilnehmern der Stichprobe 2 wurden 52 der Simulations-, 92 der Kontrollgruppe zugeteilt. Auch hier stellte sich das Studiendesign nachteilhaft für die Rekrutierung von Teilnehmern der Simulationsgruppe heraus: Insgesamt 99 Teilnehmer hatten an den Examenssimulationen teilgenommen, davon hatten 52 ebenfalls beide Umfragen beantwortet.

Unterschiedliche Stichprobenumfänge, wie hier zwischen Simulations- und Kontrollgruppe, können die Präzision des t-Tests beeinträchtigen. Da dies jedoch nicht der Fall ist, wenn die Varianzen der beiden Stichproben gleich sind [107], wurden im Vorfeld standardmäßig Levene-Tests zur Überprüfung der Varianzgleichheit durchgeführt (*siehe 3.4.2.1 Tests zur Überprüfung von Unterschiedshypothesen*).

5.1.4.2 Geschlechterverteilung

Der Anteil weiblicher Studierender an der medizinischen Fakultät der LMU lag im Sommersemester 2016 laut Pressestelle der LMU München bei 61,2%. Die Geschlechterverteilung wich in Stichprobe 1 somit gering von dieser Verteilung ab und glich ihr in Stichprobe 2 (*siehe Tabelle 3*). Ein höherer Anteil weiblicher Teilnehmer in Umfragen ist ein bereits bekanntes Phänomen, welches unter anderem von Smith [131] ausführlich diskutiert wurde.

Als Ursache kommt in diesem Zusammenhang außerdem der „Nonresponse-Bias“ in Betracht. Dieser meint Unterschiede zwischen Personen, welche eine Umfrage beantworten und jenen, die sich dagegen entscheiden. Dabei hat der Grund für die Nicht-Teilnahme eine Auswirkung auf das Ergebnis der Umfrage und kann somit deren Generalisierbarkeit beeinflussen [115, 132]. In diesem Fall wäre vorstellbar, dass sich Studierende nicht durch das Anschreiben angesprochen fühlten und somit nicht an der Umfrage teilnahmen, wenn sie nicht unter Prüfungsangst litten. Dagegen könnten vermehrt Studierende an der Umfrage teilgenommen haben, welche von Prüfungsangst betroffen waren. Wie beschrieben ist bereits bekannt, dass Frauen stärker von Prüfungsangst betroffen sind als Männer (*siehe 1.1.6 Geschlechtsspezifische Unterschiede*), der höhere Anteil weiblicher Teilnehmer könnte somit für das Vorliegen eines „Nonresponse-Bias“ sprechen. In Hinblick auf die Interpretierbarkeit und Generalisierbarkeit der Ergebnisse dieser Arbeit muss dessen potentiell Vorliegen berücksichtigt werden.

5.1.4.3 Stichprobenzuteilung

Die nicht randomisiert erfolgte Zuteilung zu Simulations- und Kontrollgruppe ist als Limitation in Hinblick auf das Studiendesign zu werten. Eine randomisierte Zuteilung wäre möglich gewesen, wenn die Anmeldezahlen die verfügbaren Plätze überstiegen hätten, oder aber die Teilnahme an der Examenssimulation nur für eine begrenzte Anzahl an interessierten Studierenden ermöglicht worden wäre. Dies war aus unserer Sicht nicht vertretbar, da durch das Sponsoring durch einen Verlag (*siehe 3.1.1.1 Durchführung von Examenssimulationen*) ausreichend Plätze für alle interessierten Studierenden verfügbar waren. In einer stark belastenden Situation, wie der Vorbereitungsphase auf das Physikum, erschien der Ausschluss von der Examenssimulation aus wissenschaftlichen Zwecken nicht angebracht.

Obwohl keine signifikanten Unterschiede zwischen Simulations- und Kontrollgruppe in den untersuchten Kategorien festgestellt werden konnten (*siehe 4.6.1 Zielgruppe der Examenssimulation*), kann das Vorliegen eines Bias durch die nicht randomisierte Zuteilung nicht ausgeschlossen werden. In den Analysen wurde dies berücksichtigt, die Stabilität (*siehe 3.4.6 Stabilität kognitiver Prüfungsangst*) wurde aus diesem Grund lediglich innerhalb der Kontrollgruppe untersucht.

5.1.5 Statistische Auswertung

Generelle Einschränkungen der statistischen Auswertung wurden bereits angesprochen: Ein Non-Response-Bias könnte vorliegen haben und sich auf die Interpretierbarkeit und Generalisierbarkeit der Ergebnisse auswirken. Zusatzmaterial und Antwortformate ermöglichen lediglich einen Überblick über Zusammenhänge und Auswirkungen bezüglich Verteilungsanalyse, Selektionsbias, Sorgen vor dem Physikum und Analyse der Examenssimulation. Auch werden Ergebnisse, welche die Examenssimulation betreffen, limitiert durch die nicht randomisiert erfolgte Zuteilung zu Simulations- und Kontrollgruppe. Zeitpunkte des mündlichen Physikums unterschieden sich zwischen den Teilnehmern stark und könnten das Antwortverhalten beeinflusst haben. Dies könnte sich auf die Verteilungsanalyse und Analyse von Ausmaß und Auslösern von Sorgen vor dem Physikum sowie nicht zuletzt auf die Teilnahme an der Examenssimulation und somit auf die damit verbundenen Analysen ausgewirkt haben. Dieser Faktor konnte in den Auswertungen nicht berücksichtigt werden, da die Termine nicht vorlagen.

Im Rahmen der Ergebnisdiskussion der statistischen Auswertung (*siehe 5.2 Diskussion der Ergebnisse*) wird auch auf mögliche Effekte und Einschränkungen durch die Methodik eingegangen. Die Methodik von Item- und Testanalyse sowie der Validierung von G-CTAS soll dagegen bereits in Teilen im folgenden Abschnitt diskutiert werden.

5.1.5.1 Item- und Testanalyse, Validierung von G-CTAS

5.1.5.1.1 Deskriptive Statistik und Itemanalyse

Empfohlene Grenzwerte für den Ausschluss von Items im Rahmen einer Itemanalyse unterscheiden sich stark [110], weshalb der Fokus auf die inhaltliche Passung des Items zur Skala gelegt wurde. Schwierigkeiten und Trennschärfeindices wurden zunächst im Verhältnis betrachtet und ein Ausschluss des Items vorgenommen, wenn es neben relativ schlechter Kennwerte außerdem inhaltlich nicht passend erschien.

5.1.5.1.2 EFA

Für die EFA wurde als Methode eine PAF mit Promax-Rotation gewählt. Russell [133] empfiehlt generell die Verwendung einer PAF, Bühner [109] die Verwendung einer PAF oder der Maximum-Likelihood-Methode (ML). Gegen die Verwendung der ML wurde entschieden, da diese vor allem sinnvoll ist, wenn im Anschluss eine konfirmatorische

Faktorenanalyse (CFA) durchgeführt wird. Da aufgrund bisheriger Untersuchungsergebnisse mit CTAS und seiner Übersetzungen von einer Korrelation der Faktoren ausgegangen werden konnte, wurde eine oblique Rotationsmethode gewählt. Darunter stellt die Promax-Rotation die Methode der Wahl dar [109], weshalb diese für die Berechnungen verwendet wurde.

Zur Bestimmung der Faktorenzahl wurden Kaiser's Kriterium und Scree-Plot beurteilt sowie eine Parallel-Analyse durchgeführt. Laut Bühner [109] handelt es sich bei der Parallelanalyse um die zuverlässigste unter diesen Methoden, dennoch lohnt auch die Betrachtung von Kaiser's Kriterium und Scree-Plot, da vor allem letztgenannter den Eigenwertverlauf anschaulich darstellt. Prinzipiell sollte für eine Faktorenanzahl entschieden werden, wenn diese zu einer plausibel zu interpretierenden Faktorenstruktur führt [109]. Hierauf wurde bei der Wahl der Faktorenanzahl letztlich der größte Wert gelegt.

Als grundlegende Einschränkung dieser Arbeit muss die Durchführung einer EFA ohne Prüfung des erarbeiteten Modells auf seine empirische Passung gesehen werden. Hierfür wäre die anschließende Durchführung einer CFA notwendig gewesen [109].

5.2 Diskussion der Ergebnisse

5.2.1 Zusammenfassung der Ergebnisse

Item- und Testanalyse, welche eine Überprüfung der Itemschwierigkeiten sowie Trennschärferechnungen umfassten, führten zum Ausschluss von insgesamt fünf Items. Mit den verbliebenen Items wurde eine exploratorische Faktorenanalyse (EFA) zur Analyse der Skalenstruktur durchgeführt. Diese zeigte eine zweidimensionale Struktur auf, welche das Vorliegen des bereits beschriebenen Methodenfaktors [75, 77] durch die Verwendung invers gepolter Items bestätigte. Aufgrund nicht ausreichend hoher Korrelation der beiden Faktoren wurden Items des zweiten Faktors von der Skala entfernt und eine eindimensionale Skala mit zwölf Items generiert (G-CTAS). Die mittlere Schwierigkeit von G-CTAS lag unter der festgelegten optimalen Schwierigkeit eines Items. Skalensummenwerte korrelierten stark und hoch signifikant mit Summenwerten der originalen 27-Item-Skala. Cronbach's alpha bestätigte die Reliabilität der Skala, durch sehr signifikante, moderate Korrelationen mit erwarteten und realen Prüfungsleistungen konnte die Inhaltsvalidität bestätigt werden.

Die Verteilungsanalyse ergab eine linksschiefe Verteilung der G-CTAS-Werte, weibliche Teilnehmer wiesen signifikant höhere Werte auf als männliche Teilnehmer.

Ein Selektionsbias zwischen Stichprobe 2 und 3 konnte bezüglich Alter und kognitiver Prüfungsangst festgestellt werden: Teilnehmer der Stichprobe 2 waren signifikant jünger und wiesen niedrigere G-CTAS-Werte auf als Teilnehmer der Stichprobe 3.

Untersuchungen zur Stabilität kognitiver Prüfungsangst ergaben keine signifikanten Unterschiede zwischen G-CTAS-Werten in Umfrage I und II, die Werte korrelierten miteinander stark und sehr signifikant.

Die Sorgen vor dem mündlichen Physikum waren hoch signifikant höher als vor dem schriftlichen Physikum. Sorgen weiblicher Teilnehmer vor dem mündlichen Physikum waren sehr signifikant höher als die der männlichen, bezüglich des schriftlichen Physikums ergaben sich keine signifikanten Unterschiede. Mögliche Zusammenhänge konnten zwischen Sorgen vor schriftlichen und mündlichen Physikum und G-CTAS-Werten festgestellt werden sowie zwischen erwarteten Prüfungsleistungen im mündlichen Physikum und Sorgen vor dem mündlichen Physikum. Bei den Analysen getrennt nach Geschlechtern ließen sich diese Zusammenhänge nur bezüglich der kognitiven Prüfungsangst unter den weiblichen Teilnehmern reproduzieren. Keine Zusammenhänge zeigten sich zwischen Prüfungssorgen und Lernaufwand, erwarteten Prüfungsleistungen im schriftlichen Physikum, realen Prüfungsleistungen oder Abiturnoten.

In den Analysen zur Zielgruppe der Examenssimulation unterschied sich die Simulationsgruppe in keinem der untersuchten Punkte von der Kontrollgruppe.

Bezüglich der Analysen möglicher Effekte durch die Teilnahme an der Examenssimulation konnte im direkten Vergleich zwischen Simulations- und Kontrollgruppe lediglich unter den männlichen Teilnehmern der Kontrollgruppe ein signifikant besseres Abschneiden im mündlichen Physikum festgestellt werden.

In den verbundenen Tests konnten keine signifikanten Unterschiede bezüglich der kognitiven Prüfungsangst, der Sorgen vor schriftlichem oder mündlichem Physikum und der Prüfungsleistung im schriftlichen Physikum festgestellt werden. Unterschiede zwischen der erwarteten und realen Prüfungsleistung im mündlichen Physikum waren dagegen teilweise signifikant: Männliche Teilnehmer der Kontrollgruppe schnitten signifikant schlechter ab als erwartet, die Kontrollgruppe hoch signifikant besser als erwartet. Der Effekt innerhalb der Kontrollgruppe war auf die weiblichen Teilnehmer zurückzuführen.

5.2.2 Item- und Testanalyse, Validierung von G-CTAS

5.2.2.1 Deskriptive Statistik und Itemanalyse, EFA

Im Rahmen der Itemanalyse wurde die Skala um mehrere Items gekürzt. Aufgrund relativ schlechter Kennwerte bezüglich Schwierigkeit und Varianz wurde Item 13 von der Skala entfernt. Mögliche Gründe hierfür wurden bereits unter Abschnitt 5.1.3.2 *Übersetzung ins Deutsche* erläutert.

Des Weiteren wurden Item 3, 10, 25 und 26 aufgrund relativ niedriger Trennschärfeindices sowie mangelhafter inhaltlicher Passung zur übrigen Skala von dieser entfernt. Item 3 („Mir fällt es weniger schwer als dem durchschnittlichen Studenten, Prüfungsanweisungen präzise aufzufassen.“) und Item 10 („Ich habe weniger Schwierigkeiten als der durchschnittliche Student, Lehrbuchinhalte zu lernen.“) wurden nach inhaltlicher Prüfung entfernt, da sie in erster Linie Lernschwierigkeiten und Textverständnis abfragen. Zwar spielt der Vergleich mit anderen Studierenden eine wichtige Rolle in der Definition von kognitiver Prüfungsangst, dennoch sind beide Items eher als mögliche, aber unspezifische Symptome oder Ursachen kognitiver Prüfungsangst aufzufassen (*siehe 1.1.4.1 Prüfungsangst und Prüfungsleistung*) und nicht in deren Definition inbegriffen.

Item 25 („Ich fühle mich ziemlich unter Druck, gute Noten zu schreiben.“) ist eher als Ausdruck von Leistungsdruck als von kognitiver Prüfungsangst zu sehen und wurde deshalb entfernt. Das Item war bereits von Furlan, Cassady & Pérez [75] als zur übrigen Skala unpassend identifiziert und deshalb nicht in CTAR aufgenommen worden.

Item 26 („Ich schneide nicht besonders gut in Prüfungen ab.“) wurde entfernt, da es als einziges Item eine objektive Frage darstellt. Dieser Schritt wurde auch in Hinblick auf eine mögliche Verfälschung der durchzuführenden Überprüfung der Inhaltsvalidität durch Korrelation mit realen Prüfungsleistungen (*siehe 3.4.3.4 Inhaltsvalidität*)

vorgenommen. Es ist aus diesem Grund kritisch zu sehen, dass das Item Bestandteil aller überarbeiteten Version von CTAS und CTAR darstellt und weiterhin bei der Überprüfung der Validität mithilfe der Prüfungsleistung berücksichtigt wird.

Die EFA ergab eine inhaltlich plausible, zweidimensionale Struktur. Alle invers gepolten Items luden auf den zweiten Faktor, womit die Ergebnisse von Furlan, Cassady & Pérez [75] sowie Cassady & Finch [77] bezüglich eines Methodenfaktors invers gepolter Items auch für die deutsche Version der Skala bestätigt werden konnten.

Item 2 („Während einer wichtigen Prüfung denke ich darüber nach, ob die anderen Studenten vielleicht besser abschneiden als ich.“) und 16 („Ich mache mir größere Gedanken darüber, in Tests gut abzuschneiden, als ich eigentlich sollte.“) luden auf keinen der beiden Faktoren ausreichend. Beide Items waren zuvor durch relativ niedrige Trennschärfeindizes aufgefallen, waren aber als inhaltlich passend eingestuft worden. Die Ergebnisse der EFA bestätigten dies somit nicht. Gemeinsam mit Item 26 thematisieren alle drei Items das Abschneiden in einer Prüfung. Die suboptimale Passung zur Skala lässt sich möglicherweise dadurch erklären, dass lediglich das Abschneiden ohne Bezug zu Angstsymptomen thematisiert wird.

Item 1 („Bevorstehende Prüfungen bereiten mir schlaflose Nächte.“) und 7 („Zu Beginn einer Prüfung bin ich so nervös, dass ich oft nicht klar denken kann.“) zeigten Querladungen. Item 1 erschien inhaltlich fraglich passend zur Skala, da es mitunter körperliche Symptome abfragt. Ein Methodeneffekt durch dessen Positionierung an erster Stelle des Fragebogens wäre als weitere Ursache für die nicht eindeutig mögliche Zuordnung zu einem Faktor denkbar.

Bei der Betrachtung des Histogramms von Item 7 fällt auf, dass dessen Verteilung stark von einer Normalverteilung abwich und kein klares Maximum zeigte (*siehe Abbildung 12*). Das Vorliegen der Querladung könnte auf die ungewöhnliche Verteilung des Items zurückzuführen sein.

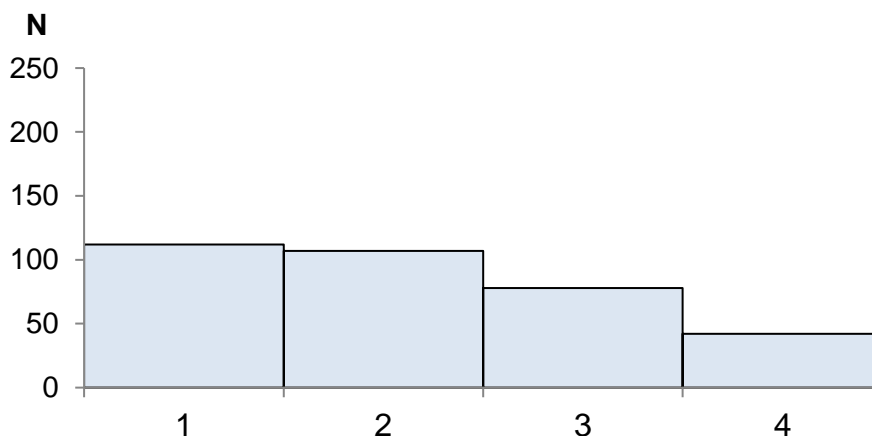


Abbildung 12 – Histogramm Item 7

Stichprobe 1, N=339.

Die Korrelation zwischen beiden Faktoren betrug 0,69. Dies zeigt zunächst, dass beide Faktoren miteinander verwandt sind. Dass der Wert nicht höher ausfällt lässt allerdings darauf schließen, dass zwischen den verschiedenen gepolten Items nicht nur methodische, sondern auch inhaltliche Unterschiede bestehen.

Alle invers gepolten Items sowie Items mit geringen Ladungen (Items 2 und 16) und Querladungen (Item 1 und 7) wurden von der Skala entfernt. G-CTAS bestand somit aus zwölf Items, welche allein und hoch auf den ersten Faktor luden.

5.2.2.2 Überprüfung der Validität von G-CTAS

Die positiven und signifikanten Korrelationen zwischen G-CTAS-Werten und erwarteten Prüfungsleistungen sowie realen Prüfungsleistungen bestätigten die Validität von G-CTAS. Im Gegensatz zur Korrelation zwischen G-CTAS I und der realen Prüfungsleistung im schriftlichen Physikum, lag die Korrelation zwischen G-CTAS I und der realen Prüfungsleistung im mündlichen Physikum knapp unter dem festgelegten Grenzwert von $r_p=0,3$. Eine mögliche Erklärung für die geringer ausfallende Korrelation ist, dass die Bewertung der mündlichen Prüfung stets abhängig vom Prüfer erfolgt. Eine mündliche Prüfung ist somit weniger standardisiert und damit subjektiver als eine schriftliche Prüfung [134]. Keine Korrelationen konnten zwischen Abiturnoten und G-CTAS-Werten festgestellt werden. Wichtig ist die Berücksichtigung der Tatsache, dass in der Regel überdurchschnittliche Abiturnoten nötig sind, um zum Medizinstudium zugelassen zu werden. Dies bestätigt ein Blick auf die Abiturnoten der Stichprobe 2 (siehe Tabelle 9). Geringe Unterschiede bezüglich der Abiturnoten der Teilnehmer erklären eine fehlende Korrelation zwischen diesen und G-CTAS-Werten.

Bei der Betrachtung der Korrelationen fällt als Nebenbeobachtung auf, dass G-CTAS II stärker mit den realen Prüfungsleistungen korrelierte als G-CTAS I. Die Prüfungsleistung könnte zu einer leichten Veränderung der kognitiven Prüfungsangst geführt haben: Die Prüfungsangst erfolgreicher Studierender nahm ab, während die weniger erfolgreicher Studierender zunahm. Dies spricht gegen die Defizithypothese und stattdessen für die Vermutung, dass Prüfungsleistungen Prüfungsangst positiv oder negativ beeinflussen können. Diese Beobachtung sollte in weiteren Untersuchungen an größeren Stichproben überprüft werden.

5.2.2.3 Vorteile und Einschränkungen bezüglich der Nutzung von G-CTAS

Die hohe Korrelation mit der 27-Items Version zeigte, dass die durch Item- und Testanalyse sowie EFA generierte 12-Item Skala (G-CTAS) nicht nur eine inhaltlich zu favorisierende, sondern außerdem eine bezüglich des Zeitaspekts ökonomisch sinnvolle Kurzform der ursprünglichen Skala darstellt.

Alle zwölf Items von G-CTAS sind in der überarbeiteten Version der englischen Originalskala (CTAR) [76] enthalten.

Die durchschnittliche Schwierigkeit der Items von G-CTAS (p_s) lag unter der optimalen Itemschwierigkeit von $p_m=2,5$, die Schwierigkeit des Fragebogens ist somit als relativ hoch einzuschätzen [110]. Die Skala eignet sich damit weniger als Screeninginstrument als vielmehr zur Differenzierung der Ausprägungsgrade kognitiver Prüfungsangst unter betroffenen Studierenden. Dennoch ist eine Modifikation der Skala durch Entfernung schwieriger Items prinzipiell möglich und würde die Skala somit auch für die Nutzung als Screeninginstrument qualifizieren. Hierfür wäre zunächst eine psychometrische Überprüfung und Validierung der modifizierten Skala nötig.

Die Reliabilität war vergleichbar mit bisherigen Untersuchungen mit CTAS und seiner Kurzformen [34, 75, 77]. Obwohl es laut Schecker [135] keinen objektiven Grenzwert von Cronbach's alpha für die Verwendung einer Skala gibt, empfehlen Autoren überwiegend einen Wert von $\geq 0,7$ [136] bzw. $\geq 0,8$ [113] anzustreben. Die Skalenreliabilität ist somit als gut einzuschätzen, sie lag außerdem knapp unter dem kritischen Wert von 0,9, welcher für eine Redundanz von Items sprechen würde [137]. Die generierte Skala bringt außerdem den Vorteil der Eindimensionalität mit sich, welcher sich vor allem in der Vereinfachung bezüglich der Bildung von und Rechnung mit Skalensummenwerten äußert.

Nach Durchführung von Test- und Itemanalyse, EFA und Validierung stellt G-CTAS eine geeignete Skala speziell zur Messung von kognitiver Prüfungsangst unter Medizinstudierenden zum Zeitpunkt des Physikums dar. Die Validierung an dieser spezialisierten Kohorte schränkt die zukünftige Nutzung des Fragebogens jedoch ein. Vor dessen Verwendung an einer von dieser Stichprobe abweichenden Population bedarf es zunächst einer erneuten Validierung.

Ebenso sollte im nächsten Schritt eine Überprüfung der empirischen Passung des Modells mittels CFA vorgenommen werden.

5.2.3 Verteilungsanalyse kognitiver Prüfungsangst, Analyse von Geschlechtsunterschieden

In der Verteilungsanalyse zeigte sich eine linksschiefe Verteilung der G-CTAS-Werte. Dies ist mitunter dadurch zu erklären, dass es sich wie unter 5.2.2.3 *Vorteile und Einschränkungen bezüglich der Nutzung von G-CTAS* beschrieben um einen schwierigen Fragebogen handelt.

Laut Powell ist die Prüfungsangstprävalenz unter Medizinstudierenden höher als unter Vergleichsgruppen [91] (*siehe 1.2.2 Medizinstudierende als Risikogruppe*). Nicht sinnvoll erschien daher die Bildung von allgemein gültigen Normwerten auf Grundlage der durchgeführten Verteilungsanalyse.

Wie erwartet waren G-CTAS-Werte der weiblichen Teilnehmer signifikant höher als die der männlichen Teilnehmer. Die Betrachtung von Cohens d zeigte, dass es sich dabei lediglich um einen kleinen Effekt handelte. Mittelwert, Modus und Median der

männlichen Teilnehmern lagen deutlich unter jenen der weiblichen Teilnehmer (*siehe Tabelle 7*), dennoch lag das Maximum bei den männlichen Teilnehmern höher, auch war die Streubreite der Werte größer. Obwohl Männer der Literatur [22, 27, 52-55, 67] sowie unserer Studie zufolge weniger stark von kognitiver Prüfungsangst betroffen sind als Frauen, fanden sich unter den männlichen Teilnehmern einige Individuen, welche auffallen hohe Summenwerte angaben. Bei einem schwierigen Fragebogen mit einem maximal zu erreichenden Summenwert von 48 sprechen Werte von 45 oder 46 für eine maximale Ausprägung kognitiver Prüfungsangst. Es handelte sich dabei nur um wenige männliche Individuen. Dies erklärt, warum es sich lediglich um einen kleinen Effekt handelte.

Die Forschung widmete Unterschieden zwischen wenig und stark betroffenen männlichen Individuen bisher kaum Aufmerksamkeit. Die Ergebnisse liefern Hinweise für Ausprägungsunterschiede kognitiver Prüfungsangst innerhalb der männlichen Teilnehmerkohorte, welche einer weiteren Abklärung bedürfen.

5.2.4 Selektionsbias

Die Überprüfung auf das Vorliegen von Unterschieden zwischen Teilnehmern einer (Stichprobe 3) und Teilnehmern beider Umfragen (Stichprobe 2) wurde in erster Linie vorgenommen, um einen Bias ausschließen zu können. Dies war Voraussetzung, um Untersuchungsergebnisse der Stichprobe 1 (bestehend aus Teilnehmern der Stichprobe 2 und 3) bzw. 2 auf die jeweils andere Stichprobe übertragen zu können.

Teilnehmer der Stichprobe 2 waren signifikant jünger und zeigten signifikant niedrigere G-CTAS-Werte als Teilnehmer der Stichprobe 3. Effektstärken waren in beiden Kategorien klein, dennoch sollen mögliche Erklärungen und Auswirkungen diskutiert werden.

Niedrigere G-CTAS-Werte könnten durch einen Non-Response-Bias verursacht worden sein: Prüfungsängstliche Studierende könnten während der Teilnahme an Umfrage I auf diese Problematik aufmerksam geworden sein und sich gegen die Teilnahme an Umfrage II entschieden haben, um einer erneuten Konfrontation zu entgehen. Wie Studien gezeigt haben, besteht ein Zusammenhang zwischen Prüfungsangst und Verzögerungsverhalten („procrastination“) [138]. In Anbetracht dessen wäre eine weitere mögliche Erklärung, dass besonders prüfungsängstliche Studierende die Teilnahme am realen Physikum auf das folgende Semester aufschoben und aufgrund dessen nicht an Umfrage II teilnahmen.

Eine vorstellbare Erklärung für das jüngere Durchschnittsalter der Teilnehmer von Umfrage II könnte sein, dass jüngere Studierende motivierter waren, an beiden Umfragen teilzunehmen als ältere Kommilitonen. Sax et al. [130] sehen die Ursache sinkender Rücklaufquoten bei Studierenden in einem Überangebot von Umfragen. Ältere Studierende waren einem Überangebot vermutlich länger ausgesetzt, sie könnten

aufgrund dessen weniger motiviert gewesen sein, auch an der zweiten Umfrage teilzunehmen.

5.2.5 Stabilität kognitiver Prüfungsangst

Wie zuvor dargestellt existieren nur wenige Untersuchungen zur Stabilität kognitiver Prüfungsangst. Cassady konnte zeigen, dass Werte im Vorfeld wichtiger Prüfungen jeweils auf einem ähnlichen Level liegen [15]. Nicht untersucht wurden dagegen individuelle Veränderungen kognitiver Prüfungsangst im Vorfeld und nach einer wichtigen Prüfung.

G-CTAS-Werte der Teilnehmer unterschieden sich zwischen den Umfragen nicht voneinander, Werte in Umfrage I korrelierten stark und hoch signifikant mit Werten in Umfrage II. Dies spricht für die individuelle Stabilität kognitiver Prüfungsangst.

Dennoch sollten bezüglich der Stabilität auch die im Abschnitt 5.2.2.2 *Überprüfung der Validität von G-CTAS* diskutierten Ergebnisse berücksichtigt werden. Diese lieferten Anhaltspunkte für einen möglichen Einfluss der Prüfungsleistung auf die kognitive Prüfungsangst.

5.2.6 Ausmaß von Sorgen vor dem Physikum und mögliche Zusammenhänge

In den Analysen wurde deutlich, dass die Sorgen vor dem mündlichen Physikum sehr viel größer waren als vor dem schriftlichen. Über 60% der Teilnehmer stimmten der Aussage vollständig zu, dass das mündliche Physikum ihnen Sorgen bereite.

Korrelationen zeigten, dass Sorgen vor schriftlichem und mündlichem Physikum teilweise durch die kognitive Prüfungsangst, Sorgen vor dem mündlichen Physikum außerdem durch erwartete Prüfungsleistung im mündlichen Physikum zu erklären waren. Die Korrelationen lagen nur knapp über dem festgelegten Grenzwert von $r_p=0,3$, die Zusammenhänge sind demzufolge als schwach einzustufen. Unter den weiblichen Teilnehmern ließen sich diese nur bezüglich der kognitiven Prüfungsangst, unter den männlichen Teilnehmern in keiner Kategorie reproduzieren. Zwischen Prüfungssorgen und erwarteter Prüfungsleistung im schriftlichen Physikum, Lernaufwand oder Abiturnote konnten keine Zusammenhänge festgestellt werden.

Das Physikum stellt somit eine Prüfung dar, welche zu einem großen Teil unabhängig von primär in Frage kommenden Faktoren Sorgen auslöste.

Sorgen sind definiert als „zukunftsorientierter Gefühlszustand, in welchem man sich auf den Versuch vorbereitet, mit einem bevorstehenden negativen Event umzugehen“ [139]. Sie sind laut Holodynski [1] und Liebert & Morris [7] Teil der Definition von Prüfungsangst. Es macht deshalb Sinn, sich bei der Suche alternativer Auslöser der Prüfungssorgen auch Ursachen von Prüfungsangst ins Gedächtnis zu rufen (*siehe 1.1.2 Entstehung*).

Erschwerte Lernbedingungen, Erwartungen des Umfelds sowie Ungewissheit bezüglich Prüfungen und ihren Anforderungen wurden hier als wichtige Punkte genannt. In Anbetracht der unter *1.2.1 Der Erste Abschnitt der Ärztlichen Prüfung (Physikum)* geschilderten Vorgaben und Abläufe der Prüfung wird deutlich, inwiefern diese speziell im Rahmen des primär betroffenen mündlichen Physikums eine Rolle spielen könnten: Termine variieren an der LMU München im Gegensatz zum schriftlichen Physikum stark und werden teilweise erst sehr kurzfristig veröffentlicht. Der Prüfungsstoff unterscheidet sich durch die verschiedenen Prüfer oft auch innerhalb eines Fachgebiets stark, auch wird die genaue Prüferkombination erst wenige Tage im Vorfeld der Prüfung bekanntgegeben. Die mündliche Prüfung wird außerdem gemeinsam mit Kommilitonen abgehalten und ist dadurch weniger anonym als die schriftliche Prüfung. Neben der Tatsache, dass Studierende schriftliche Prüfungen als fairer einschätzen als mündliche [140], könnten all diese Faktoren erklären, warum das mündliche Physikum zu größeren Sorgen führte als das schriftliche. Nicht zu vernachlässigen ist außerdem die Tatsache, dass es sich beim Physikum aufgrund der Relevanz für die Gesamtnote und die begrenzten Wiederholungsmöglichkeiten um eine wichtige Prüfung des Medizinstudiums handelt (*siehe 1.2.1 Der Erste Abschnitt der Ärztlichen Prüfung (Physikum)*).

Unsere Ergebnisse decken sich mit den Beobachtungen von Núñez-Peña, Suárez-Pellicioni et al. [57]: Frauen hatten signifikant mehr Sorgen vor der mündlichen Prüfung als Männer. Bezüglich der schriftlichen Multiple-Choice-Prüfung unterschieden sich die Sorgen dagegen nicht.

Durch den Vergleich mit den realen Prüfungsleistungen im Physikum wurde deutlich, dass die Sorgen nicht mit diesen in Zusammenhang standen. Somit handelte es sich um irrationale Angst. In Hinblick auf die Arbeit von Pintado und Llamazares [39] ergibt sich hieraus eine Konsequenz für mögliche Interventionen: Psychologische Interventionen scheinen primär indiziert, Angebote zur Vermittlung von Lern- und Vorbereitungsmethoden hingegen sekundär.

5.2.7 Examenssimulationen

5.2.7.1 Zielgruppe der Examenssimulation

Untersucht wurden Unterschiede zwischen Simulations- und Kontrollgruppe bezüglich Prüfungsangst und Prüfungssorgen, Lernverhalten sowie bisherigen und erwarteten Prüfungsleistungen. Dies war zum einen nötig, um einen Selektionsbias zwischen den Gruppen ausschließen zu können, zum anderen um die von der Examenssimulation angesprochene Zielgruppe näher zu untersuchen und die Ergebnisse des Formats auf seine Nützlichkeit zu überprüfen.

Simulations- und Kontrollgruppe unterschieden sich in den untersuchten Punkten nicht voneinander.

Bezüglich der Auswertung der Effekte der Examenssimulation brachte das Fehlen signifikanter Unterschiede in den untersuchten Kategorien für den Vergleich zwischen Simulations- und Kontrollgruppe Vorteile mit sich, da ähnliche Ausgangspunkte vorlagen.

Das Fehlen signifikanter Unterschiede zeigt, dass das Angebot eine heterogene Zielgruppe ansprach und nicht besonders prüfungsängstliche, leistungsschwache oder bezüglich Lernaufwand wenig motivierte Studierende. Dies hätte den Vorteil mit sich gebracht, diese während der belastenden Situation durch den direkten Kontakt leichter unterstützen zu können.

Im Gegenzug ist positiv zu werten, dass es sich bei Examenssimulationen offenbar um ein Prüfungsvorbereitungsformat handelte, welches unter einer großen Bandbreite an Studierenden Anklang fand.

5.2.7.2 Effekte der Examenssimulation

5.2.7.2.1 Kognitive Prüfungsangst

Die kognitive Prüfungsangst unterschied sich zwischen Simulations- und Kontrollgruppe nach dem Physikum nicht signifikant voneinander. Auch bezüglich der Unterschiede von individuellen G-CTAS-Werten in Umfrage I und II in den verbundenen Tests konnten keine signifikanten Unterschiede festgestellt werden.

Im Abschnitt 5.2.5 *Stabilität kognitiver Prüfungsangst* wurde bereits diskutiert, dass es sich bei kognitiver Prüfungsangst um ein stabiles Konstrukt zu handeln scheint. Dass es weder bei der Simulations- noch bei der Kontrollgruppe zu einer signifikanten Veränderung der Werte kam, spricht für diese These.

Ebenfalls vorstellbar ist, dass die einmalige Durchführung der Examenssimulation nicht ausreichend war, um kognitive Prüfungsangst positiv zu beeinflussen. Dies könnte mithilfe eines veränderten Studiendesigns überprüft werden. So könnte die wiederholte Teilnahme an Examenssimulationen, möglicherweise bereits seit Beginn bzw. vor der eigentlichen Vorbereitungszeit auf die Prüfung, einen sinnvollen nächsten Schritt darstellen.

5.2.7.2.2 Sorgen vor schriftlichem und mündlichem Physikum

Auch die retrospektiven Sorgen unterschieden sich zwischen Simulations- und Kontrollgruppe nicht signifikant voneinander. Ebenfalls waren weder in Simulations- noch in Kontrollgruppe signifikante Unterschiede von individuellen Sorgen in Umfrage I und II festzustellen.

Die Examenssimulation sollte über eine ermöglichte realistische Einschätzung des Lernstands und der voraussichtlichen Prüfungsleistung zur Reduktion von Angst und

Sorgen vor der bevorstehenden Prüfung führen (*siehe 1.1.5.1 Prüfungssimulationen*). Wie unter 5.2.6 *Ausmaß von Sorgen vor dem Physikum und mögliche Zusammenhänge* dargestellt, waren die Prüfungssorgen jedoch irrational. Lernstand und erwartete Prüfungsleistung im schriftlichen Physikum waren nicht ausschlaggebend für die Entstehung der Sorgen, das Format könnte demzufolge ungeeignet gewesen sein, Sorgen zu reduzieren.

Über Konfrontation und Desensibilisierung bieten Prüfungssimulationen die Möglichkeit, auch gegen irrationale Angst vorzugehen. Da Sorgen vor dem mündlichen Physikum signifikant höher waren als vor dem schriftlichen und hier signifikante Korrelationen mit der erwarteten Prüfungsleistung vorlagen (*siehe Tabelle 12*), sollten in weiteren Schritten Simulationen des mündlichen Physikums auf deren Potential zur Reduktion von Angst und Sorgen vor einer Prüfung überprüft werden.

5.2.7.2.3 Lernaufwand

Auch bei der Analyse des Lernaufwands waren keine signifikanten Unterschiede zwischen Simulations- und Kontrollgruppe zu verzeichnen.

Falls nötig sollte die Examenssimulation durch eine realistische Einschätzung des Lernstands zu einer Intensivierung des Lernaufwands führen. Die Mittelwerte in der Kategorie „Lernaufwand“ (*siehe Tabelle 9 und 13*) zeigen, dass dieser bereits in Umfrage I unter allen Teilnehmern hoch war, egal ob diese an der Examenssimulation teilnahmen oder nicht. Bedingt durch die hohen Ausgangswerte konnte es nicht mehr zu wesentlichen Steigerungen kommen, was sich mit Blick auf Mittelwerte der Umfrage II bestätigt (*siehe Tabelle 14*). Somit ist auch die fehlende Ausbildung von Unterschieden zwischen Simulations- und Kontrollgruppe nicht verwunderlich.

Um die Auswirkungen von Prüfungssimulationen auf den Lernaufwand zu überprüfen, sollten zukünftige Untersuchungen die Intensität der Prüfung variieren und den Lernaufwand mit alternativen Methoden messen.

5.2.7.2.4 Prüfungsleistungen

Bezüglich des schriftlichen Physikums unterschieden sich Prüfungsleistungen zwischen Simulations- und Kontrollgruppe nicht signifikant voneinander. Weder in der Simulations- noch in der Kontrollgruppe konnten signifikante Unterschiede zwischen individuellen erwarteten und realen Prüfungsleistungen im schriftlichen Physikum festgestellt werden. Die Teilnahme an der Examenssimulation führte somit zu keiner verbesserten Prüfungsleistung der Teilnehmer im schriftlichen Physikum.

Die hohe Übereinstimmung von erwarteten und realen Prüfungsleistungen im schriftlichen Physikum zeigt, dass die Selbsteinschätzung der Studierenden bereits zum Zeitpunkt von Umfrage I sehr gut war. Studierende konnten ihre Prüfungsleistungen bereits vor Durchführung der Simulation realistisch abschätzen. Dies könnte im Umkehrschluss bedeuten, dass die Zeit zwischen Simulation und realem schriftlichen

Physikum zu kurz war, um die Vorbereitung anzupassen und die in der Simulation identifizierten Wissenslücken zu schließen. Für Folgestudien ergibt sich daraus, dass die Durchführung von Prüfungssimulationen weiter im Vorfeld der realen Prüfung geeigneter sein könnte, die Prüfungsleistung der Teilnehmer zu verbessern.

Eine weitere Erklärung für die fehlenden Unterschiede bezüglich der realen Prüfungsleistung im schriftlichen Physikum bietet auch der Lernaufwand auf das schriftliche Physikum: Werte zwischen Simulations- und Kontrollgruppe unterschieden sich nicht voneinander. Demzufolge betrieben Studierende, welche nicht an der Examenssimulation teilnahmen, einen ebenso hohen Lernaufwand wie Teilnehmer der Examenssimulation. In Hinblick auf die Prüfungsleistung stellte sich dieser als nicht weniger effektiv heraus.

Eine Möglichkeit zur Weiterentwicklung des Formats würde hier die Verwendung von freien Fragen anstatt Multiple-Choice-Fragen darstellen. Diese bringen einen höheren Lerneffekt mit sich [141, 142] und könnten so die Prüfungsleistung der Teilnehmer der Examenssimulation verbessern.

Männliche Teilnehmer der Kontrollgruppe schnitten im mündlichen Physikum signifikant besser ab als Teilnehmer der Simulationsgruppe. Bei den weiblichen Teilnehmern zeigten sich bezüglich der realen Prüfungsleistungen im mündlichen Physikum keine signifikanten Unterschiede zwischen Simulations- und Kontrollgruppe.

Der mögliche Schluss, dass die Teilnahme an der Examenssimulation zu einer schlechteren Prüfungsleistung der männlichen Teilnehmer im mündlichen Physikum führte, greift hier vermutlich zu kurz. Eine andere Erklärung wäre stattdessen, dass Studierende, welche sich gegen eine Teilnahme entschieden, den Fokus in ihrer Vorbereitung auf das mündliche Physikum legten und dies durch bessere reale Prüfungsleistungen belohnt wurde.

Die Tatsache, dass männliche Teilnehmer der Simulationsgruppe signifikant schlechter im mündlichen Physikum abschnitten als erwartet, fügt sich in dieses Bild. Diese Teilnehmer rechneten mit besseren Leistungen im mündlichen Physikum und legten möglicherweise deshalb den Fokus auf die Vorbereitung auf das schriftliche Physikum. Dies machte sich umgekehrt mit schlechteren realen Prüfungsleistungen im mündlichen Physikum bezahlt.

Teilnehmer der Kontrollgruppe schnitten im mündlichen Physikum dagegen signifikant besser ab als erwartet. Auch dies ist durch einen gesteigerten Vorbereitungsaufwand bei geringerer erwarteter Prüfungsleistung denkbar. Der Effekt war dabei vor allem auf die weiblichen Teilnehmer zurückzuführen. Dass weibliche Teilnehmer dazu neigen, sich bezüglich Prüfungsleistungen schlechter einzuschätzen, steht im Einklang mit bisherigen Forschungsergebnissen [143] und bietet eine weitere Erklärung für die

aufgezeigten Unterschiede zwischen erwarteter und realer Prüfungsleistung der weiblichen Teilnehmer der Kontrollgruppe im mündlichen Physikum.

6 Fazit & Ausblick

1. Übersetzung der „Cognitive Test Anxiety Scale“ ins Deutsche und deren psychometrische Validierung an Medizinstudierenden vor dem Physikum.

Übersetzung und psychometrische Validierung der Skala konnten erfolgreich vorgenommen werden. G-CTAS stellt ein eindimensionales, reliables und valides Messinstrument kognitiver Prüfungsangst dar. Methodische und inhaltliche Effekte durch die Verwendung invers kodierter Items konnten repliziert werden. Obwohl mittlerweile eine überarbeitete englischsprachige Version vorliegt, ist eine Anpassung von G-CTAS nicht notwendig, da alle Items der deutschen Version in der neuen Version enthalten sind.

Die Verwendung der 12-Items Skala eignet sich aufgrund ihrer Schwierigkeit primär nicht als Screening-Instrument.

Die Zielsetzung spricht direkt einen relativen Nachteil der Skala an: Diese ist zur Messung an Medizinstudierenden vor dem Physikum validiert worden, vor Verwendung an einer abweichenden Population sollte eine erneute Überprüfung der Skala vorgenommen werden. Dies wäre auch in Hinblick auf die Etablierung von Normwerten sinnvoll. Ebenso sollte die Passung des Konstrukts mittels CFA in naher Zukunft überprüft werden.

Daten einer größeren Kohorte Studierender verschiedener Fakultäten der LMU München liegen hierfür bereits vor, Ergebnisse stehen zum aktuellen Zeitpunkt jedoch noch aus.

2. Verteilungsanalyse der Skalensummenwerte unter Medizinstudierenden unmittelbar vor dem Physikum.

G-CTAS-Werte zeigten eine linksschiefe Verteilung, was unter anderem auf die Schwierigkeit des Fragebogens zurückzuführen ist. Wie erwartet war die kognitive Prüfungsangst unter den weiblichen Teilnehmern signifikant höher als unter den männlichen Teilnehmern. Die Verteilung der Werte unterschied sich zwischen den Geschlechtern, dies sollte an größeren Stichproben weiter untersucht werden. Auffällig war die teilweise extrem hohe Ausprägung kognitiver Prüfungsangst unter einzelnen männlichen Individuen, was ebenfalls in Zukunft näher untersucht werden sollte.

3. Überprüfung der Stabilität kognitiver Prüfungsangst durch Vergleich von Werten vor und nach dem Physikum.

Untersuchungsergebnisse sprachen einerseits für die Stabilität kognitiver Prüfungsangst, zum anderen lieferten sie Anhaltspunkte für den potentiellen Einfluss der Prüfungsleistung auf die kognitive Prüfungsangst. Auch diese Ergebnisse erfordern weiteren Abklärungsbedarf.

4. Bestimmung des Ausmaßes an Prüfungssorgen und Analyse von Zusammenhängen zum besseren Verständnis erhöhter Stressbelastung durch das Physikum.

Die mündliche Physikumsprüfung konnte als stark belastende Prüfung identifiziert werden. Lediglich schwache Zusammenhänge konnten zwischen Prüfungssorgen und kognitiver Prüfungsangst sowie erwarteten Prüfungsleistungen im mündlichen Physikum in Bezug auf das mündliche Physikum festgestellt werden.

Als weitere Ursachen kommen vor allem in Hinblick auf das mündliche Physikum Unsicherheiten bezüglich der Organisation der Prüfung, der Prüfer und des Prüfungsstoffs sowie die Prüfungsbewertung in Betracht. Um dies näher zu untersuchen, wurden im vergangenen Herbst in einer groß angelegten deutschlandweiten Studie Daten gesammelt. Neben Prüfungsdaten wurden außerdem Resilienzmessungen vorgenommen sowie Burnout- und Stresssymptome quantifiziert. Ergebnisse stehen zum aktuellen Zeitpunkt noch aus.

Die Ergebnisse dieser Arbeit legen nahe, dass nicht nur ein hoher Bedarf für Vorbereitungsangebote auf das schriftliche, sondern auch auf das mündliche Physikum besteht. Da die Sorgen unabhängig von der realen Prüfungsleistung und somit häufig irrational waren, sollten neben der inhaltlichen Vorbereitung auf die Prüfung auch Coping-Strategien vermittelt und psychologisch orientierte Interventionen angeboten werden. Unsere Erfahrungen haben gezeigt, dass bei einigen Studierenden der Bedarf hierfür besteht. Psychologische Interventionen stellen allerdings sicher nicht für alle Studierende ein geeignetes Instrument dar.

Trotz der für gewöhnlich überdurchschnittlichen schulischen Leistungen Medizinstudierender kann eine Prüfung wie das Physikum eine enorme Belastung darstellen. Obwohl es in der Literatur zahlreiche Hinweise auf erhöhte Stressbelastung und Burnout-Gefahr unter Medizinstudierenden und Ärzten gibt (*siehe 1.2.2 Medizinstudierende als Risikogruppe*), ist der Einfluss von Prüfungen und speziell des Physikums kaum untersucht. Wir sehen hier enormen Nachholbedarf, da diese in Anbetracht unserer Ergebnisse als (Mit-)Auslöser durchaus in Frage kommen. Ein erster Schritt in diese Richtung ist bereits mit der oben genannten Studie unternommen worden.

5. Durchführung von Prüfungssimulationen des schriftlichen Physikums sowie Untersuchung der Zielgruppe und möglicher Effekte auf kognitive Prüfungsangst, Prüfungssorgen, Lernaufwand und Prüfungsleistungen

Die Examenssimulationen konnten wiederholt erfolgreich durchgeführt werden. Es handelte sich nicht nur um ein sehr gern genutztes Angebot mit hohen Teilnehmerzahlen, die Eignung als Prüfungsvorbereitungsangebot bestätigte sich auch durch die heterogene Zielgruppe.

Generell wäre bei erneuter wissenschaftlicher Testung des Formats eine randomisierte Zuteilung in Simulations- und Kontrollgruppe sinnvoll.

Die Teilnahme an der Examenssimulation konnte weder kognitive Prüfungsangst oder Sorgen vor den Prüfungen reduzieren, noch konnte der Lernaufwand gesteigert oder die Prüfungsleistungen der Teilnehmer verbessert werden. Gründe hierfür könnten unter anderem die Stabilität kognitiver Prüfungsangst, die Irrationalität der Prüfungssorgen, der generell hohe Vorbereitungsaufwand und der Durchführungszeitpunkt darstellen.

Für die Zukunft kämen eine häufigere Durchführungssequenz der Examenssimulationen mit der Möglichkeit der wiederholten Teilnahme, die Durchführung weiter im Vorfeld der realen Prüfung oder die Verwendung eines abgewandelten Fragenformats in Betracht.

In Hinblick auf *4.5 Ausmaß von Sorgen vor dem Physikum und mögliche Zusammenhänge* sollten auch Prüfungsvorbereitungsangebote auf das mündliche Physikum erprobt werden, wie zum Beispiel die Durchführung mündlicher Examenssimulationen. Simulationen und Peer-Teaching-Formate des dritten Abschnitts der Ärztlichen Prüfung werden an der LMU München bereits angeboten und deren Potential in Rahmen wissenschaftlicher Arbeiten untersucht. Teilnehmerzahlen und vorläufige Analysen scheinen vielversprechend, endgültige Ergebnisse stehen zum aktuellen Zeitpunkt jedoch ebenfalls noch aus.

7 Zusammenfassung

Vor allem die kognitive Prüfungsangst steht aufgrund des nachgewiesenen Zusammenhangs mit verminderter Prüfungsleistung im Fokus der Forschung. Im deutschen Sprachraum ist bislang kein geeignetes Messinstrument für deren Erfassung verfügbar, wenig untersucht sind außerdem Verteilung und Stabilität individueller kognitiver Prüfungsangst im Kontext einer wichtigen Prüfung.

Unter Medizinstudierenden könnten unter anderem Prüfungen für die wiederholt nachgewiesene erhöhte Stressbelastung und Angstprävalenz verantwortlich sein. Nicht untersucht ist, inwieweit eine relevante Prüfung wie das Physikum Sorgen bei den Studierenden auslösen kann und mit welchen Faktoren diese korreliert sind.

Prüfungssimulationen könnten ein geeignetes Format zur Reduktion von kognitiver Prüfungsangst und Prüfungssorgen sowie zur Steigerung von Lernaufwand und Prüfungsleistungen darstellen.

Die Studie wurde von Herbst 2015 bis Herbst 2016 im Vorher-Nachher-Vergleich durch Befragung von Studierenden wenige Wochen vor und nach dem schriftlichen Physikum durchgeführt. Vor der realen schriftlichen Prüfung wurden freiwillige Simulationen des schriftlichen Physikums angeboten.

Die im Rahmen dieser Arbeit übersetzte und psychometrisch validierte Version der „Cognitive Test Anxiety Scale“ [67] (G-CTAS) erwies sich als eindimensionales, reliables und valides Messinstrument kognitiver Prüfungsangst. Werte kognitiver Prüfungsangst zeigten eine linksschiefe Verteilung und unterschieden sich zwischen den Geschlechtern. Die Stabilität konnte nicht endgültig bestätigt werden, es ergaben sich außerdem Hinweise auf einen möglichen Einfluss von Prüfungsleistungen auf die kognitive Prüfungsangst.

Das mündliche Physikum stellte sich im Vergleich mit dem schriftlichen als die stärker belastende Prüfung heraus. Es handelte sich dabei um irrationale Angst, welche durch Unsicherheiten in Bezug auf die Prüfung verursacht sein könnte. Psychologisch orientierte Interventionen sollten deshalb Teil des Prüfungsvorbereitungsangebots werden, dieses sollte außerdem auf das mündliche Physikum ausgeweitet werden.

Bei den Examenssimulationen des schriftlichen Physikums handelte es sich um ein sehr beliebtes Angebot mit heterogener Zielgruppe. Die Teilnahme konnte weder zu einer Reduktion kognitiver Prüfungsangst oder von Prüfungssorgen führen, noch konnte der Lernaufwand gesteigert oder die Prüfungsleistungen verbessert werden. Die Stabilität kognitiver Prüfungsangst, die Irrationalität der Prüfungssorgen, der generell hohe Vorbereitungsaufwand und der Durchführungszeitpunkt könnten mögliche Gründe hierfür darstellen.

Literaturverzeichnis

1. Holodynski, M. (2019). *Prüfungsangst*. In M. A. Wirtz (Hrsg.), *Dorsch – Lexikon der Psychologie*. Abgerufen am 17.05.2019, von <https://m.portal.hogrefe.com/dorsch/pruefungsangst-1/>
2. Diagnostic, A. P. A. (2017). *Statistical Manual of Mental Disorders (DSM–5)* American Psychiatric Association website 2017 Available from: <https://www.psychiatry.org/psychiatrists/practice/dsm>, zuletzt abgerufen am 17.05.2019
3. *CD-10-GM Version 2019, Systematisches Verzeichnis, Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision, Stand: 21.September 2018*. 2018, Köln: Deutsches Institut für Medizinische Dokumentation und Information (DIMDI) im Auftrag des Bundesministeriums für Gesundheit (BMG) unter Beteiligung der Arbeitsgruppe ICD des Kuratoriums für Fragen der Klassifikation im Gesundheitswesen (KKG).
4. Dachs, L., *Prüfungsängste*, in *Lehrbuch der Verhaltenstherapie, Band 3: Psychologische Therapie bei Indikationen im Kindes- und Jugendalter*, S. Schneider and J. Margraf, Editors. 2019, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 569-586.
5. Metzigg, W. and M. Schuster, *Prüfungsangst und Lampenfieber als Bewertungsangst*, in *Prüfungsangst und Lampenfieber*. 2018, Springer. p. 3-7.
6. Fehm, L. and T. Fydrich, *Prüfungsangst*. Vol. 44. 2011: Hogrefe Verlag.
7. Liebert, R. and L. Morris, *Cognitive and emotional components of test anxiety: A distinction and some initial data*. Psychological Reports, 1967. **20**: p. 975-978.
8. Sarason, I.G., *Stress, anxiety, and cognitive interference: Reactions to tests*. Journal of Personality and Social Psychology, 1984. **46(4)**: p. 929-938.
9. Hodapp, V., *Das Prüfungsängstlichkeitsinventar TAI-G: Eine erweiterte und modifizierte Version mit vier Komponenten [The Test Anxiety Inventory TAI-G: An expanded and modified version with four components]*. Zeitschrift für Pädagogische Psychologie / German Journal of Educational Psychology, 1991. **5(2)**: p. 121-130.
10. Pekrun, R., et al., *Beyond test anxiety: Development and validation of the test emotions questionnaire (TEQ)*. Anxiety, Stress & Coping, 2004. **17(3)**: p. 287-316.
11. Pekrun, R. and T. Götz, *Emotionsregulation: Vom Umgang mit Prüfungsangst*. Handbuch Lernstrategien, 2006: p. 248-258.
12. Putwain, D.W. and B. Aveyard, *Is perceived control a critical factor in understanding the negative relationship between cognitive test anxiety and examination performance?* School Psychology Quarterly, 2018. **33(1)**: p. 65-74.

13. Reeve, C.L., S. Bonaccio, and J.E. Charles, *A policy-capturing study of the contextual antecedents of test anxiety*. Personality and Individual Differences, 2008. **45**(3): p. 243-248.
14. Lotz, C. and J.R. Sparfeldt, *Does test anxiety increase as the exam draws near?—Students' state test anxiety recorded over the course of one semester*. Personality and Individual Differences, 2017. **104**: p. 397-400.
15. Cassady, J.C., *The stability of undergraduate students' cognitive test anxiety levels*. Practical Assessment, Research & Evaluation, 2001. **7** (20).
16. Spielberger, C.D., *Anxiety: Current trends in theory and research*. Elsevier. 1972.
17. Spielberger, C.D., W.D. Anton, and J. Bedell, *The nature and treatment of test anxiety*. Emotions and anxiety: New concepts, methods, and applications. 2015. 317-344.
18. Steinmayr, R., et al., *Subjective Well-Being, Test Anxiety, Academic Achievement: Testing for Reciprocal Effects*. Frontiers in Psychology, 2016. **6**.
19. Pixner, S. and L. Kaufmann, *Prüfungsangst, Schulleistung und Lebensqualität bei Schülern*. Lernen und Lernstörungen, 2013.
20. Schaefer, A., et al., *Mental health and performance of medical students with high and low test anxiety*. Psychotherapie, Psychosomatik, medizinische Psychologie, 2007. **57**(7): p. 289-297.
21. Akinsola, E.F. and A.D. Nwajei, *Test anxiety, depression and academic performance: assessment and management using relaxation and cognitive restructuring techniques*. Psychology, 2013. **4**(06): p. 18.
22. Hembree, R., *Correlates, causes, and treatment of test anxiety*. Review of Educational Research, 1988. **58**: p. 47-77.
23. Morris, L.W. and R.M. Liebert, *Effects of anxiety on timed and untimed intelligence tests: Another look*. Journal of Consulting and Clinical Psychology, 1969. **33**(2): p. 240-244.
24. Morris, L.W., M.A. Davis, and C.H. Hutchings, *Cognitive and emotional components of anxiety: Literature review and a revised worry-emotionality scale*. Journal of Educational Psychology, 1981. **74**(4): p. 541-555.
25. Deffenbacher, J.L. and S.L. Hazaleus, *Cognitive, emotional, and physiological components of Test Anxiety*. Cognitive Therapy and Research, 1985. **9**(2): p. 169-180.
26. Thomas, C.L., J.C. Cassady, and M.L. Heller, *The influence of emotional intelligence, cognitive test anxiety, and coping strategies on undergraduate academic performance*. Learning and Individual Differences, 2017. **55**: p. 40-48.

27. Putwain, D.W., *Test anxiety and GCSE performance: The effect of gender and socio - economic background*. Educational Psychology in Practice, 2008. **24**(4): p. 319-334.
28. Covington, M.V. and C.L. Omelich, "*I knew it cold before the exam*": A test of the *anxiety-blockage hypothesis*. Journal of educational psychology, 1987. **79**(4): p. 393.
29. Wine, J., *Test anxiety and direction of attention*. Psychological Bulletin 1971. **76**(2): p. 92-104.
30. Reeve, C.L. and S. Bonaccio, *Does test anxiety induce measurement bias in cognitive ability tests?* Intelligence, 2008. **36**(6): p. 526-538.
31. Sommer, M. and M.E. Arendasy, *Comparing different explanations of the effect of test anxiety on respondents' test scores*. Intelligence, 2014. **42**: p. 115-127.
32. Sommer, M. and M.E. Arendasy, *Further evidence for the deficit account of the test anxiety–test performance relationship from a high-stakes admission testing setting*. Intelligence, 2015. **53**: p. 72-80.
33. Cassady, J.C., *Anxiety in schools: The causes, consequences, and solutions for academic anxieties*. Vol. 2. 2010: Peter Lang.
34. Cassady, J.C., *The influence of cognitive test anxiety across the learning–testing cycle*. Learning and Instruction, 2004. **14**(6): p. 569-592.
35. Pekrun, R., *Expectancy-value theory of anxiety: Overview and implications*. 1992.
36. von der Embse, N., J. Barterian, and N. Segool, *Test Anxiety Interventions for Children and Adolescents: A Systematic Review of Treatment Studies from 2000-2010*. Psychology in the Schools, 2013. **50**(1): p. 57-71.
37. Ergene, T., *Effective interventions on test anxiety reduction: A meta-analysis*. School Psychology International, 2003. **24**(3): p. 313-328.
38. Tryon, G.S., *The Measurement and Treatment of Test Anxiety*. Review of Educational Research, 1980. **50**(2): p. 343-372.
39. Pintado, I.S. and M.d.C.E. Llamazares, *Description of the General Procedure of a Stress Inoculation Program to Cope with the Test Anxiety*. Psychology, 2014. **05**(08): p. 956-965.
40. Daly, A.L., S. Chamberlain, and V. Spalding, *Test anxiety, heart rate and performance in A-level French speaking mock exams: an exploratory study*. Educational Research, 2011. **53**(3): p. 321-330.
41. Young, I., et al., *The benefits of a peer - assisted mock OSCE*. The clinical teacher, 2014. **11**(3): p. 214-218.

42. Smeds, M.R., et al., *Use of mock oral examinations in vascular surgery training programs: a nationwide survey*. Journal of Surgical Research, 2018. **232**: p. 94-98.
43. Roth, C., et al., *Therapie von Angst- und Zwangspatienten in der verhaltenstherapeutischen Praxis*. Verhaltenstherapie, 2004. **14**(1): p. 16-21.
44. Christmann, F., *Keine Angst vor Ängsten: Verhaltenstherapeutische Techniken lernen und anleiten*. 2018: Klett-Cotta.
45. Maercker, A. and A. Weike, *Systematische Desensibilisierung*, in *Lehrbuch der Verhaltenstherapie, Band 1*. 2018, Springer. p. 403-409.
46. Warnecke, I., *Prüfungsangst bewältigen: Ein Trainingsprogramm in sieben Schritten*. Vol. 4857. 2017: UTB.
47. Bensberg, G. and J. Messer, *Was tun bei Prüfungsangst?*, in *Survivalguide Bachelor*. 2010, Springer. p. 209-225.
48. Saravanan, C. and R. Kingston, *A randomized control study of psychological intervention to reduce anxiety, amotivation and psychological distress among medical students*. Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences, 2014. **19**(5): p. 391.
49. Larsen, D.P., A.C. Butler, and H.L. Roediger III, *Test - enhanced learning in medical education*. Medical education, 2008. **42**(10): p. 959-966.
50. Larsen, D.P., A.C. Butler, and H.L. Roediger III, *Comparative effects of test - enhanced learning and self - explanation on long - term retention*. Medical Education, 2013. **47**(7): p. 674-682.
51. McDaniel, M.A., et al., *Testing the testing effect in the classroom*. European Journal of Cognitive Psychology, 2007. **19**(4-5): p. 494-513.
52. Zeidner, M., *Does test anxiety bias scholastic aptitude test performance by gender and sociocultural group?* J Pers Assess, 1990. **55**(1-2): p. 145-60.
53. Chapell, M.S., et al., *Test Anxiety and Academic Performance in Undergraduate and Graduate Students*. Journal of Educational Psychology, 2005. **97**(2): p. 268-274.
54. Nilofer Farooqi, Y., R. Ghani, and C. D. Spielberger, *Gender Differences in Test Anxiety and Academic Performance of Medical Students*. International Journal of Psychology and Behavioral Sciences, 2012. **2**(2): p. 38-43.
55. Lowe, P.A., *Should Test Anxiety be Measured Differently for Males and Females? Examination of Measurement Bias Across Gender on Measures of Test Anxiety for Middle and High School, and College Students*. Journal of Psychoeducational Assessment, 2015. **33**(3): p. 238-246.

56. Sharma, S. and A. Sud, *Examination Stress and Test Anxiety: A Cross-Cultural Perspective*. Psychology & Developing Societies, 1990. **2(2)**: p. 183-201.
57. Núñez-Peña, M.I., M. Suárez-Pellicioni, and R. Bono, *Gender Differences in Test Anxiety and Their Impact on Higher Education Students' Academic Achievement*. Procedia - Social and Behavioral Sciences, 2016. **228**: p. 154-160.
58. Spielberger, C.D., *Test Anxiety Inventory: Preliminary Professional Manual*. . Palo Alto, CA: Consulting Psychologists Press, 1980.
59. Hodapp, V., *The TAI-G: A multidimensional approach to the assessment of test anxiety*. Stress, anxiety, and coping in academic settings, 1996: p. 95-130.
60. Hodapp, V., S. Rohrmann, and T. Ringeisen, *Prüfungsangstfragebogen (PAF) [German Test Anxiety Questionnaire]*. Göttingen, Germany: Hogrefe, 2011.
61. Keith, N., et al., *Cross-sectional and longitudinal confirmatory factor models for the german test anxiety inventory: A construct validation*. Anxiety, Stress & Coping, 2003. **16(3)**: p. 251-270.
62. Ringeisen, T., P. Buchwald, and V. Hodapp, *Capturing the multidimensionality of test anxiety in cross-cultural research: An English adaptation of the German Test Anxiety Inventory*. Cognition, Brain, Behavior: An Interdisciplinary Journal, 2010. **14(4)**: p. 347-364.
63. Musch, J. and A. Bröder, *Psychometrische Eigenschaften und Validität des multidimensionalen Prüfungsängstlichkeitsinventars TAI-G*. Zeitschrift für Pädagogische Psychologie, 1999. **13(1/2)**: p. 100-105.
64. Benson, J. and D. L. Bandalos, *Second-Order Confirmatory Factor Analysis of the Reactions to Tests Scale with Cross-Validation*. Vol. 27. 1992. 459-487.
65. Quast, H.H., M. Jerusalem, and J.G. Sarason, *Reaktions-Tendenzen gegenüber Tests RTT-G (Reactions to Tests)*, in *Institute for Psychology*, R. Schwarzer, Editor. 1986, Free University of Berlin: Skalen zur Befindlichkeit und Persönlichkeit p. 139-155.
66. Quast, H.H., *Reactions to Tests*, in *Departement of Psychology - Educational Psychology*., R. Schwarzer, Editor. 1985, Berlin: Freie Universität: Stress and social support. p. 255-258.
67. Cassady, J.C. and R.E. Johnson, *Cognitive Test Anxiety and Academic Performance*. Contemporary Educational Psychology, 2002. **27(2)**: p. 270-295.
68. Ramirez, G. and S.L. Beilock, *Writing About Testing Worries Boosts Exam Performance in the Classroom*. Science, 2011. **331**: p. 211-213.
69. Cassady, J.C., *The impact of cognitive test anxiety on text comprehension and recall in the absence of external evaluative pressure*. Applied Cognitive Psychology, 2004. **18(3)**: p. 311-325.

70. Tsianos, N., et al., *An experimental assessment of the use of cognitive and affective factors in adaptive educational hypermedia*. IEEE Transactions on Learning Technologies, 2009. **2**: p. 249-258.
71. Kapetanaki, E.M. *Test anxiety in education: Policies and practices*. in *Presentation at the 31st World Conference on Stress and Anxiety Research*. 2010, August 4-6. Galway, Ireland.
72. Cassady, J.C., A. Mohammed, and L. Mathieu, *Cross-Cultural Differences in Test Perceptions*. Journal of Cross-Cultural Psychology, 2004. **35**(6): p. 713-718.
73. Chen, M.L., *Test anxiety, reading anxiety, and reading performance among university English as a second language learners*. 2007, Ming Chuan University, Taipei, Taiwan.
74. Zheng, Y., *Chinese University Students' Motivation, Anxiety, Global Awareness, Linguistic Confidence, and English Test Performance: A Correlational and Causal Investigation*. 2010.
75. Furlan, L.A., J.C. Cassady, and E.R. Pérez, *Adapting the Cognitive Test Anxiety Scale for use with Argentinean University Students*. International Journal of Testing, 2009. **9**(1): p. 3-19.
76. Cassady, J.C. and W.H. Finch, *Using factor mixture modeling to identify dimensions of cognitive test anxiety*. Learning and Individual Differences, 2015. **41**: p. 14-20.
77. Cassady, J.C. and W.H. Finch, *Confirming the Factor Structure of the Cognitive Test Anxiety Scale: Comparing the Utility of Three Solutions*. Educational Assessment, 2014. **19**(3): p. 229-242.
78. Thomas, C.L., J.C. Cassady, and W.H. Finch, *Identifying Severity Standards on the Cognitive Test Anxiety Scale: Cut Score Determination Using Latent Class and Cluster Analysis*. Journal of Psychoeducational Assessment, 2018. **36**(5): p. 492-508.
79. Baghaei, P. and J. Cassady, *Validation of the Persian Translation of the Cognitive Test Anxiety Scale*. SAGE Open, 2014. **4**(4): p. 215824401455511.
80. Bozkurt, S., et al., *Validation of the Turkish Version of the Cognitive Test Anxiety Scale–Revised*. SAGE Open, 2017. **7**(1): p. 215824401666954.
81. Approbationsordnung für Ärzte vom 27. Juni 2002 (BGBl. I S. 2405), die zuletzt durch Artikel 5 des Gesetzes vom 17. Juli 2017 (BGBl. I S. 2581) geändert worden ist.
82. München, LMU, *Medizin 1. Studienabschnitt (Staatsexamen)*. https://www.uni-muenchen.de/studium/studienangebot/studiengaenge/studienfaecher/medizin_1/staatsexam/index.html, zuletzt abgerufen am 01.02.2019, Stand: 21.12.2018.

83. München, LMU, *Erster Abschnitt der Ärztlichen Prüfung*. https://www.uni-muenchen.de/studium/administratives/pruefungsaemter/07_med/pruefungsamt_mediz/index.html, verantwortlich für den Inhalt: Ref. III.6, zuletzt abgerufen am 05.05.2019.
84. Herbst, U., et al., *Studierendenstress in Deutschland – eine empirische Untersuchung*. AOK-Bundesverband, 2016.
85. Dyrbye, L.N., M.R. Thomas, and T.D. Shanafelt, *Systematic review of depression, anxiety, and other indicators of psychological distress among U.S. and Canadian medical students*. Acad Med, 2006. **81**(4): p. 354-73.
86. Rotenstein, L.S., et al., *Prevalence of Depression, Depressive Symptoms, and Suicidal Ideation Among Medical Students: A Systematic Review and Meta-Analysis*. JAMA, 2016. **316**(21): p. 2214-2236.
87. Dahlin, M., N. Joneborg, and B. Runeson, *Stress and depression among medical students: A cross - sectional study*. Medical education, 2005. **39**(6): p. 594-604.
88. Jurkat, H.B., et al., *Depressivität und Stressbewältigung bei Medizinstudierenden. Eine Vergleichsuntersuchung des 1. und 7. Fachsemesters Humanmedizin*. Nervenarzt, 2011. **82**(5): p. 646-52.
89. Iqbal, S., S. Gupta, and E. Venkatarao, *Stress, anxiety & depression among medical undergraduate students & their socio-demographic correlates*. The Indian journal of medical research, 2015. **141**(3): p. 354.
90. Puthran, R., et al., *Prevalence of depression amongst medical students: a meta - analysis*. Medical education, 2016. **50**(4): p. 456-468.
91. Powell, D.H., *Behavioral treatment of debilitating test anxiety among medical students*. Journal of Clinical Psychology, 2004. **60**(8): p. 853-865.
92. Burger, P., et al., *Vom Studienstart bis zum ersten Staatsexamen – Zunahme von Depressivität bei gleichzeitigem Verlust des Kohärenzgefühls und der psychischen Lebensqualität in höheren Semestern Humanmedizin*. Vol. 64. 2014. 322-327.
93. Moffat, K.J., et al., *First year medical student stress and coping in a problem - based learning medical curriculum*. Medical education, 2004. **38**(5): p. 482-491.
94. Ludwig, A.B., et al., *Depression and stress amongst undergraduate medical students*. BMC Med Educ, 2015. **15**: p. 141.
95. Willcock, S.M., et al., *Burnout and psychiatric morbidity in new medical graduates*. Med J Aust, 2004. **181**(7): p. 357-60.
96. Voltmer, E.K., Thomas; Westermann, Jürgen, *Prävention: Gesund durchs Medizinstudium*. Deutsches Ärzteblatt, 2015. **112**: p. 35-36.

97. Dyrbye, L. and T. Shanafelt, *A narrative review on burnout experienced by medical students and residents*. Medical education, 2016. **50**(1): p. 132-149.
98. Panagioti, M., et al., *Association Between Physician Burnout and Patient Safety, Professionalism, and Patient Satisfaction: A Systematic Review and Meta-analysis*. JAMA Intern Med, 2018. **178**(10): p. 1317-1330.
99. Bugaj, T.J., et al., *Stresserleben bei Medizinstudierenden: Welche Rolle spielen Bindungserleben und Persönlichkeitsaspekte?* PPM-Psychotherapie-Psychosomatik· Medizinische Psychologie, 2016. **66**(02): p. 88-92.
100. Kurth, R.A., et al., *Studienbezogene Belastungen, Lebensqualität und Beziehungserleben bei Medizinstudenten*. Psychotherapeut, 2007. **52**(5): p. 355-361.
101. Rosenthal, J.M. and S. Okie, *White coat, mood indigo—depression in medical school*. New England journal of medicine, 2005. **353**(11): p. 1085-1088.
102. Dreimüller, N., *Medizinstudenten sind häufiger depressiv*. InFo Neurologie & Psychiatrie, 2017. **19**(3): p. 22-22.
103. Pramataroff-Hamburger, V., et al., *Erfahrungen aus einer Balintgruppe für Medizinstudenten*. Balint Journal, 2013. **14**(04): p. 115-120.
104. Horn, J.L., *A rationale and test for the number of factors in factor analysis*. Psychometrika, 1965. **30**(2): p. 179-185.
105. Patil, V.H., et al., *Parallel Analysis Engine to Aid Determining Number of Factors to Retain* [Computer software] (2007). Abgerufen unter: <http://smishra.faculty.ku.edu/parallelengine.htm>, zuletzt abgerufen am 02.04.2019.
106. Lenhard, W. and A. Lenhard, *Berechnung von Effektstärken*. Abgerufen unter: <https://www.psychometrica.de/effektstaerke.html>. Dettelbach: Psychometrica. DOI: 10.13140/RG.2.1.3478.4245, zuletzt abgerufen am 02.04.2019, 2016.
107. Bortz, J. and C. Schuster, *Statistik für Human- und Sozialwissenschaftler. Lehrbuch mit Online-Materialien*. 2011: Springer Berlin Heidelberg.
108. Cohen, J., *Statistical power analysis for the behavioral sciences*. 1988: Routledge.
109. Bühner, M., *Einführung in die Test- und Fragebogenkonstruktion*, ed. a.u.e.A. 3. 2011, München: Pearson Studium. 640.
110. Mummendey, H.D. and I. Grau, *Die Fragebogen-Methode: Grundlagen und Anwendung in Persönlichkeits-, Einstellungs- und Selbstkonzeptforschung*. 2014: Hogrefe Verlag.
111. Kaiser, H.F., *The Application of Electronic Computers to Factor Analysis*. Educational and Psychological Measurement, 1960. **20**(1): p. 141-151.

112. Cattell, R.B., *The Scree Test For The Number Of Factors*. Multivariate Behavioral Research, 1966. **1**(2): p. 245-276.
113. Bortz, J. and N. Döring, *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. 2006, Heidelberg: Springer.
114. Zwick, W.R. and W.F. Velicer, *Comparison of five rules for determining the number of components to retain*. Psychological Bulletin, 1986. **99**(3): p. 432-442.
115. Phillips, A.W., S. Reddy, and S.J. Durning, *Improving response rates and evaluating nonresponse bias in surveys: AMEE Guide No. 102*. Med Teach, 2016. **38**(3): p. 217-28.
116. Maurer, M. and O. Jandura, *Masse statt Klasse? Einige kritische Anmerkungen zu Repräsentativität und Validität von Online-Befragungen.*, in *Sozialforschung im Internet*, N. Jakob, H. Schoen, and T. Zerback, Editors. 2009: VS Verlag für Sozialwissenschaften, Wiesbaden.
117. Handwerk, P.G., C. Carson, and K.M. Blackwell, *On-Line vs. Paper-and-Pencil Surveying of Students: A Case Study*. AIR 2000 Annual Forum Paper. 2000.
118. Underwood, D., H. Kim, and M. Matier, *To Mail or To Web: Comparisons of Survey Response Rates and Respondent Characteristics*. 2000: AIR 2000 Annual Forum Paper.
119. Szafranski, D.D., T.L. Barrera, and P.J. Norton, *Test anxiety inventory: 30 years later*. Anxiety, Stress, & Coping, 2012. **25**(6): p. 667-677.
120. Rost, D. and F. Schermer, «*Reaktionsweisen gegenüber Tests*» (RTT) und «*Manifestationen von Leistungsangst*» (DAI-MAN): *una eademque res?* [*"Reactions to Tests" (RTT) and "Manifestations of Test Anxiety" (DAI-MAN): Una Eademque Res?*]. Zeitschrift für Differentielle und Diagnostische Psychologie, 1989. **10**: p. 169-179.
121. Hughes, G.D., *The Impact of Incorrect Responses to Reverse-Coded Survey Items*. Research in the Schools, 2009. **16**(2).
122. Nardi, P.M., *Doing Survey Research : A Guide to Quantitative Methods*. 2003, Boston: Allyn and Bacon.
123. Rotthoff, T., et al., *Assessing the learning environment of a faculty: psychometric validation of the German version of the Dundee Ready Education Environment Measure with students and teachers*. Med Teach, 2011. **33**(11): p. e624-36.
124. Busch, A.-K., et al., *Do medical students like communication? Validation of the German CSAS (Communication Skills Attitude Scale)*. GMS Zeitschrift für Medizinische Ausbildung, 2015. **32.1 (2015): Doc11**.
125. Moosbrugger, H. and A. Kelava, *Testtheorie und Fragebogenkonstruktion*. 2 ed. 2012: Springer-Verlag Berlin Heidelberg.

126. Institut für medizinische und pharmazeutische Prüfungsfragen, *Archiv Medizin* <https://www.impp.de/pruefungen/medizin/archiv-medizin.html>, zuletzt abgerufen am 30.04.2019.
127. Sheehan, K.B., *E-mail Survey Response Rates: a Review*. Journal of Computer-Mediated Communication, 2001. **6**(2): p. JCMC621.
128. De Leeuw, E. and W. Heer, *Trends in household survey nonresponse: A longitudinal and international comparison.*, in *Survey Nonresponse*, R.M. Groves, et al., Editors. 2002, John Wiley & Sons, New York. p. 41-54.
129. Manfreda, K.L., et al., *Web Surveys versus other Survey Modes: A Meta-Analysis Comparing Response Rates*. International Journal of Market Research, 2008. **50**(1): p. 79-104.
130. Sax, L.J., S.K. Gilmartin, and A.N. Bryant, *Assessing Response Rates and Nonresponse Bias in Web and Paper Surveys*. Research in Higher Education, 2003. **44**:409.
131. Smith, G., *Does gender influence online survey participation?: A record-linkage analysis of university faculty online survey response behavior*. ERIC Document Reproduction Service No. ED 501717, 2008.
132. Groves, R.M., *Nonresponse Rates and Nonresponse Bias in Household Surveys*. Public Opinion Quarterly, 2006. **70**(5): p. 646-675.
133. Russell, D.W., *In search of underlying dimensions: The use (and abuse) of factor analysis in Personality and Social Psychology Bulletin*. Personality and social psychology bulletin, 2002. **28**(12): p. 1629-1646.
134. Campbell, F. and J. Westwood, *Oral versus written assessments: a test of student performance and attitudes AU - Huxham, Mark*. Assessment & Evaluation in Higher Education, 2012. **37**(1): p. 125-136.
135. Schecker, H., *Überprüfung der Konsistenz von Itemgruppen mit Cronbachs alpha*, in *Methoden in der naturwissenschaftsdidaktischen Forschung*. 2014, Springer-Verlag, Springer Spektrum.
136. Taber, K.S., *The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education*. Research in Science Education, 2018. **48**(6): p. 1273-1296.
137. Streiner, D.L., *Starting at the beginning: an introduction to coefficient alpha and internal consistency*. Journal of personality assessment, 2003. **80**(1): p. 99-103.
138. Custer, N., *Test anxiety and academic procrastination among prelicensure nursing students*. Nursing education perspectives, 2018. **39**(3): p. 162-163.
139. Brown, T.A., T.A. O'Leary, and D.H. Barlow, *Generalized anxiety disorder*. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders: A step-by-step treatment manual*, : New York, NY, US: The Guilford Press, 2001: p. 154-208.

140. Ehlers, J.P., *Blended Assessment: Mündliche und elektronische Prüfungen im klinischen Kontext*. Zeitschrift für Hochschulentwicklung, 2009.
141. McConnell, M.M., C. St-Onge, and M.E. Young, *The benefits of testing for learning on later performance*. Advances in Health Sciences Education, 2015. **20**(2): p. 305-320.
142. McDaniel, M.A., H.L. Roediger, and K.B. McDermott, *Generalizing test-enhanced learning from the laboratory to the classroom*. Psychonomic bulletin & review, 2007. **14**(2): p. 200-206.
143. Blanch-Hartigan, D., *Medical students' self-assessment of performance: results from three meta-analyses*. Patient education and counseling, 2011. **84**(1): p. 3-9.

Abbildungsverzeichnis

Abbildung 1 - Studiendesign	- 14 -
Abbildung 2 - Stichprobenzuteilung.....	- 21 -
Abbildung 3 - Itemschwierigkeiten.....	- 29 -
Abbildung 4 - Trennschärfeindices.....	- 31 -
Abbildung 5 – Scree-Plot	- 32 -
Abbildung 6 - Verteilung von G-CTAS-Werten	- 36 -
Abbildung 7 – Verteilung von G-CTAS-Werten weiblicher Teilnehmer	- 37 -
Abbildung 8 – Verteilung von G-CTAS-Werten männlicher Teilnehmer	- 37 -
Abbildung 9 – Histogramm „Das schriftliche Physikum bereitet mir Sorgen“	- 42 -
Abbildung 10 – Histogramm „Das mündliche Physikum bereitet mir Sorgen“	- 42 -
Abbildung 11 – Teilnehmerzahlen Umfrage I und II	- 59 -
Abbildung 12 – Histogramm Item 7.....	- 65 -
Abbildung 13 – Histogramm Item 4.....	- 113 -
Abbildung 14 – Histogramm Item 11	- 113 -
Abbildung 15 – Histogramm Item 12	- 114 -
Abbildung 16 – Histogramm Item 13 ¹	- 114 -
Abbildung 17 – Histogramm Item 17 ¹	- 115 -
Abbildung 18 – Histogramm Item 20.....	- 115 -
Abbildung 19 – Histogramm Item 23.....	- 116 -
Abbildung 20 – Histogramm Item 24.....	- 116 -

Tabellenverzeichnis

<i>Tabelle 1 - Inhalt Umfrage I: Kategorie mit zugehörigen Fragen/Messinstrumenten und Antwortskalen</i>	<i>18 -</i>
<i>Tabelle 2 - Inhalt Umfrage II: Kategorie mit zugehörigen Fragen/Messinstrumenten und Antwortskalen</i>	<i>19 -</i>
<i>Tabelle 3 - Stichprobenbeschreibungen.....</i>	<i>21 -</i>
<i>Tabelle 4 - Deskriptive Statistik der deutschen CTAS-Items</i>	<i>28 -</i>
<i>Tabelle 5 - Trennschärfeindices.....</i>	<i>30 -</i>
<i>Tabelle 6 - Exploratorische Faktorenanalyse</i>	<i>33 -</i>
<i>Tabelle 7 - Überprüfung der Inhaltsvalidität: Korrelationen zwischen G-CTAS-Werten mit erwarteten und realen Prüfungsleistungen.....</i>	<i>35 -</i>
<i>Tabelle 8 - Deskriptive Statistiken von G-CTAS-Werten</i>	<i>36 -</i>
<i>Tabelle 9 – Vergleich von Stichprobe 2 und 3 in Umfrage I.....</i>	<i>39 -</i>
<i>Tabelle 10 - Mittelwertunterschiede von G-CTAS I und II.....</i>	<i>40 -</i>
<i>Tabelle 11 – Ausmaß von Sorgen vor schriftlichem und mündlichem Physikum im Vergleich.....</i>	<i>41 -</i>
<i>Tabelle 12 – Korrelationen mit Sorgen vor schriftlichem und mündlichem Physikum- 44 -</i>	
<i>-</i>	
<i>Tabelle 13 - Vergleich von Simulations- und Kontrollgruppe in Umfrage I.....</i>	<i>46 -</i>
<i>Tabelle 14 - Vergleich von Simulations- und Kontrollgruppe in Umfrage II.....</i>	<i>48 -</i>
<i>Tabelle 15 – Unterschiede von G-CTAS-Werten in Umfrage I und II.....</i>	<i>49 -</i>
<i>Tabelle 16 - Unterschiede von Sorgen vor schriftlichem und mündlichem Physikum in Umfrage I und II.....</i>	<i>50 -</i>
<i>Tabelle 17 – Unterschiede von erwarteten Prüfungsleistungen und realen Prüfungsleistungen im schriftlichen und mündlichen Physikum.....</i>	<i>52 -</i>
<i>Tabelle 18 - Berechnungsprozess der Trennschärfeindices.....</i>	<i>117 -</i>

Abkürzungsverzeichnis

Δ	Mittlere Differenz
CFA	Konfirmatorische Faktorenanalyse
CTAR	Cognitive Test Anxiety Scale – revised
CTAS	Cognitive Test Anxiety Scale
d	Cohens d
DSM-V	„Diagnostic and Statistical Manual of Mental Disorders, 5th Edition“ [2]
EFA	Exploratorische Faktorenanalyse
G-CTAS	Cognitive Test Anxiety Scale – German version
G-CTAS I	G-CTAS-Werte in Umfrage I
G-CTAS II	G-CTAS-Werte in Umfrage II
ICD 10	„International Statistical Classification of Diseases and Related Health Problems, 10th Revision“ [3]
IMPP	Institut für medizinische und pharmazeutische Prüfungsfragen
KMO	Kaiser-Meyer-Olkin-Koeffizient
Lernaufwand I	Lernaufwand auf das Physikum in Umfrage I
Lernaufwand II	Lernaufwand (retrospektiv) auf das Physikum in Umfrage II
LMU	Ludwig-Maximilians-Universität München
M	Mittelwert
Max.	Maximum
Min.	Minimum
ML	Maximum Likelihood
MSA	Measure of Sample Adequacy
N	Anzahl
p_s	Durchschnittliche Schwierigkeit der Items
p_m	Schwierigkeitsgrad eines Items
PAF	principal axis factoring (dt. Hauptachsenanalyse)
r_{it}	Part-whole-korrigierter Diskriminationsindex
r_p	Pearsons Rangkorrelationskoeffizient
r_s	Spearman's Rangkorrelationskoeffizient
RTT	„Reactions to Tests“

SD	Standardabweichung
Sorgen I	Sorgen vor dem Physikum in Umfrage I
Sorgen II	Sorgen vor dem Physikum (retrospektiv) in Umfrage II
TAI	„Test Anxiety Inventory“ (dt. „Prüfungsangstinventar“)
TAI-G	„Test Anxiety Inventory“ – German version
V	Varianz
χ^2	Chi-Quadrat

Anhang

Antragsformular



Antragsformular Version 1.0.

Seite 1 von 4

Antragsformular für Nicht-AMG-Studien

(Ergänzend ist ein eigenständiges Studienprotokoll vorzulegen)

(s. auch „Erläuterungen des Antrags für Nicht- AMG/Nicht MPG- Studien)

1. Antragssteller: Name, Anschrift, Tel. /Fax- Nr.:	LMU-StaR-Team Leitung: PD Dr. med. Matthias Angstwurm Medizinische Klinik Innenstadt Pettenkoferstr. 8a 80336 München Tel.: +49-89-5160-52472 Fax.: +49-89-5160-2341
2. Titel des Forschungs- vorhabens:	„Strategiekonzept zur Identifizierung von Prüfungsangst und deren erfolgreicher Bewältigung im Examenkontext der Ärztlichen Prüfung“
3. Ausbildungsdaten und Prüferfahrung:	Geburtsdatum: 26.05.1962 April 1993: Approbation 06.03.2001: Facharztprüfung Internist <ul style="list-style-type: none">• Klinischer Oberarzt der internistischen Intensivstation in der Medizinischen Klinik Innenstadt• Weiterbildung zum Master of Medical Education (MME)• PJ-Beauftragter der Medizinischen Fakultät der LMU• Verantwortliche Leitung des LMU-Staatsexamens-Repetitoriums (LMU-StaR)• Erfahrung bei der Durchführung von klinischen Studien als Studienleiter (SIC-Studie), Koordinator, sowie als Prüfarzt im Zentrum (MAXSEPT, Kybersept-Studie, TNF-Antikörper bei Sepsis, mehrere Antibiotika-Studien)
4. Monozentrische oder multizentrische Studie:	monozentrische Studie <input checked="" type="checkbox"/> multizentrische Studie <input type="checkbox"/>
5. Erklärung, dass die Grundsätze der Deklaration von Helsinki in Ihrer derzeit gültigen Fassung berücksichtigt werden.	Ich erkläre, dass die Grundsätze der Deklaration von Helsinki in der derzeit gültigen Fassung berücksichtigt werden. Ort, Datum, Unterschrift
6. Studienbedingte Strahlenbelastung:	Ja <input type="checkbox"/> Nein <input checked="" type="checkbox"/>

Das Klinikum der Universität München ist eine Anstalt des öffentlichen Rechts

Großhadern: Marchioninstr. 15 - D-81337 München - Tel. +49 (0)89 4400 - 0
öffentl. Verkehr: U6, 56, 266, 269 oder N41 bis Haltestelle Klinikum Großhadern

Innenstadt: Ziemssenstr. 1 - D-80336 München - Tel. +49 (0)89 4400 - 0
öffentl. Verkehr: U1, U2, U3, U6, 16, 17, 18, N17, N27 bis Haltestelle Sendlinger Tor

7. Wissenschaftliche Angaben zum Forschungsvorhaben:	
7.1. Studienziel/Fragestellung	<p>Das medizinische Staatsexamens-Repetitorium der Ludwig-Maximilians-Universität München (LMU-StaR) ist ein wichtiger didaktischer Eckpfeiler in der Vorbereitung der Studierenden auf die Ärztliche Prüfung. Dabei bereitet das LMU-StaR-Konzept die Medizinstudierenden der LMU strukturiert und ergebnisorientiert auf die Prüfungssituation des Staatsexamens vor. Zugleich ist eine gezielte Prüfungssimulation zur Stressbewältigung der Examenssituation bisher an der LMU nicht ausreichend etabliert. Gerade der individuelle Erfolgsdruck, der empirischen Studien zufolge im Studienfach Humanmedizin besteht, führt mit der häufig aufgebauten Konkurrenzsituation zu einem negativen Reinforcement und folglich zu einer Verschlechterung der Prüfungsergebnisse. Ziel dieser Studie ist daher zunächst die Prävalenz der Prüfungsangst unter den Studierenden der Medizinischen Fakultät der LMU München durch eine strukturierte Datenanalyse zu untersuchen. Hierzu werden personenbezogene Daten (Demographie, Semesterzahl, Ergebnis der Ärztlichen Prüfung und Prüfungserfahrung, etc.) sowie persönliche Einschätzungen zur Prüfungsvorbereitung und -bewältigung (nach Cassidy-Johnson) erhoben.</p> <p>Einen weiteren Schwerpunkt der Studie wird ein Kurskonzept bilden, das den Einfluss von Prüfungssimulationen in Kleingruppen (max. 15 Personen) auf die Bewältigung der Examenssituation analysiert.</p> <p><u>Fragestellungen (siehe Studienprotokoll):</u> Besteht ein Zusammenhang zwischen der Prävalenz von Prüfungsangst und den Ergebnissen der Studierenden in der Ärztlichen Prüfung? Verändert die Teilnahme an Examenssimulationskursen den Ausprägungsgrad der Prüfungsangst respektive das Abschneiden der Studierenden in der Ärztlichen Prüfung? Lassen sich durch eine gezielte Intervention im Rahmen der Prüfungsvorbereitung individuelle Defizite definieren und reduzieren?</p>
7.2. Studiendesign	<p>Mit Kontrollgruppe Ja <input checked="" type="checkbox"/> Nein <input type="checkbox"/></p> <p>Mit Gesunden Ja <input checked="" type="checkbox"/> Nein <input type="checkbox"/></p>
7.3. Statistische Beratung - Fallzahl-schätzung und Auswertungskonzept	<p>Ja <input type="checkbox"/> Durch wen?:</p> <p>Nein <input checked="" type="checkbox"/> Bitte begründen: Steht noch aus.</p>
7.4. Studiendauer	<p>Voraussichtlicher Beginn: April 2015</p> <p>Voraussichtliches Ende: November 2015</p>

8. Diskussion der ethisch-rechtlich relevanten Probleme	
8.1. Diskussion der Nachteile, Risiken, Belastungen und des möglichen Nutzens der Studie	<p>Um ein Strategiekonzept zur Identifizierung und Bewältigung von Prüfungsangst entwickeln zu können, müssen individuelle Daten (z.B. bisherige Fach-/Semesteranzahl, Alter, Geschlecht etc.) sowie Angaben zur Prüfungsvorbereitung und Selbsteinschätzung (nach Cassady Johnson) erhoben und ausgewertet werden. Die Studierenden tragen im Rahmen einer Online-Befragung freiwillig die entsprechenden Angaben ein und erteilen ihre Einwilligung zur wissenschaftlichen Auswertung dieser Daten. Erteilt der Studierende seine Einwilligung nicht, ist die Teilnahme an den Prüfungssimulationen dennoch möglich. Die komplette Auswertung und Publikation der Daten der Studierenden, die eingewilligt haben, erfolgt in pseudonymisierter Form, so dass keinem der Studierenden durch die wissenschaftliche Auswertung persönliche Nachteile entstehen können.</p> <p>Eine Weitergabe individualisierter Daten ist nicht vorgesehen. Die erhofften Erkenntnisse können zur Qualitätssicherung des Kurskonzeptes und zur individuellen Verbesserung der Examensvorbereitung der Teilnehmer herangezogen werden. Die Studienergebnisse unterstützen die Studierenden in den verschiedenen Phasen der Examensvorbereitung und ermöglichen eine Personalisierung des Kurskonzeptes auf das individuelle Prüfungsverhalten der Studierenden.</p> <p>Durch die wiederholte Überprüfung ihres Lernfortschritts und ihres Prüfungsverhaltens können die Studierenden motiviert werden, kontinuierlich zu lernen und ihr Wissen zu verfestigen. Insofern erfolgt auch für die künftige Planung des LMU-StaR eine inhaltliche und strukturelle Weiterentwicklung.</p>
8.2. Einschluss nicht-Einwilligungsfähiger	<p>Vorgesehen <input type="checkbox"/></p> <p>Nicht vorgesehen <input checked="" type="checkbox"/></p>
8.3. Forschung an Minderjährigen	<p>Vorgesehen <input type="checkbox"/></p> <p>Nicht vorgesehen <input checked="" type="checkbox"/></p>
9. Datenschutz:	<p>Irreversibel anonymisierte Verwendung der Proben /Daten (ohne Personenbezug) <input type="checkbox"/></p> <p>Pseudonymisierte Verwendung der Proben /Daten (Verschlüsselung) <input checked="" type="checkbox"/></p> <p>Angaben, wie der Verschlüsselungscode gebildet wird, wer Zugang zum Verschlüsselungscode hat, Gründe für die Entschlüsselung und wo/wie lange die Daten aufbewahrt werden:</p> <p>Alle personenbezogenen Daten (siehe oben) werden von den Studierenden freiwillig zur Verfügung gestellt. Die Datensammlung benutzt als Kennung die Matrikelnummer. Jeder Studierende erklärt mit der Teilnahme an der Online-Befragung (via Anklicken eines Links) sein Einverständnis, dass er mit der Auswertung seiner Daten im Rahmen einer wissenschaftlichen Studie einverstanden ist. Unmittelbar nach Abschluss der Datensammlung werden alle Daten irreversibel pseudonymisiert.</p>

10. Versicherung	<p>Erklärung, dass eine Versicherung abgeschlossen wird /wurde <input type="checkbox"/></p> <p>Warum eine Versicherung nicht notwendig ist <input checked="" type="checkbox"/></p> <p>Die Studienteilnahme bringt keinerlei gesundheitliche Risiken mit sich. Es erfolgen keine invasiven Maßnahmen. Die Probanden absolvieren lediglich eine schriftliche Online-Prüfung am PC. Bei den Studienteilnehmern handelt es sich ausschließlich um immatrikulierte Medizinstudierende des vor-/klinischen Studienabschnitts.</p> <p>Wurde eine Wegeversicherung abgeschlossen</p> <p>Ja <input type="checkbox"/> Nein <input checked="" type="checkbox"/></p>
11. Datenschutz:	<p>Durch die Studie entstehen keine Zusatzkosten.</p>
12. Aufwandsentschädigung für die EK: Kostenübernahmeerklärung/Antrag auf Erlass der Aufwandsentschädigung	<p>Die Kosten der Aufwandsentschädigung werden übernommen</p> <p>Antrag auf Erlass der Aufwandsentschädigung mit Begründung <input checked="" type="checkbox"/></p> <p>Die Studie wird durch keine Mittel unterstützt, sodass ein Antrag auf Erlass der Aufwandsentschädigung gestellt wird.</p>
<p>Datum, Unterschrift</p>	



Studienprotokoll

Titel der Studie: „Strategiekonzept zur Identifizierung von Prüfungsangst und deren erfolgreicher Bewältigung im Examenskontext der Ärztlichen Prüfung“

Studienprotokoll

(Stand: 09.03.2015, Version 1.0.2015)

Verantwortlichkeiten:

PD Dr. med. Matthias Angstwurm, MME (Doktorvater, Studienleiter)

Medizinische Klinik IV

Klinikum der Universität München

Tel: +49/89/4400-52207

E-Mail: Matthias.Angstwurm@med.uni-muenchen.de

Dr. med. Dipl.-Jur. Univ. Christina Berchtold (Betreuerin)

LMU-Staatsexamens-Repetitorium

Pettenkoferstraße 8a

Klinikum der Universität München

Tel: +49/89/4400-52472

E-Mail: Christina.Berchtold@med.uni-muenchen.de

Alexandra Stefan (Doktorandin)

Medizinstudentin der LMU

Tel: +49/89/4400-52472

E-Mail: Alexandra.Stefan@med.uni-muenchen.de

Das Klinikum der Universität München ist eine Anstalt des öffentlichen Rechts

Großhadern: Marchioninstr. 15 · D-81337 München · Tel. +49 (0)89 4400 - 0
öffentl. Verkehr: U6, 56, 266, 269 oder N41 bis Haltestelle Klinikum Großhadern

Innenstadt: Ziemssenstr. 1 · D-80336 München · Tel. +49 (0)89 4400 - 0
öffentl. Verkehr: U1, U2, U3, U6, 16, 17, 18, N17, N27 bis Haltestelle Sendlinger Tor

Unterschriften zur Bestätigung des Protokolls:

<hr/> <i>Datum, Unterschrift</i> PD Dr. Matthias Angstwurm <i>Medizinische Klinik IV</i> <i>Studienleiter</i>	
<hr/> <i>Datum, Unterschrift</i> Dr. Christina Berchtold	<hr/> <i>Datum, Unterschrift</i> Alexandra Stefan

1. Rationale**1.1. Hintergrund und Stand der Forschung**

Das medizinische Staatsexamensrepetitorium der Ludwig-Maximilians-Universität München (LMU-StaR) ist ein wichtiger didaktischer Eckpfeiler in der Vorbereitung der Studierenden auf die Ärztliche Prüfung. Das Ziel des LMU-StaR-Konzeptes besteht darin, die Medizinstudierenden der LMU München strukturiert und ergebnisorientiert auf die Prüfungssituation des Staatsexamens vorzubereiten. Dabei basiert das Repetitoriums-Curriculum auf mehreren Säulen, die zugleich die Schwerpunkte des Examens repräsentieren. Bislang lag der Fokus des Kursangebotes auf der gezielten Vermittlung und Wiederholung von prüfungsrelevantem Wissen. Damit soll ein verbessertes Abschneiden der Studierenden im Staatsexamen erreicht werden, um die LMU München auch in landesweiten Rankings weiter als Spitzenuniversität positionieren zu können. Zugleich ist eine gezielte Prüfungssimulation zur Stressbewältigung der Examenssituation bisher an der LMU nicht ausreichend etabliert und stellt auch international eine innovative und interprofessionelle Kernstrategie mit hohem Verbesserungs- und Entwicklungspotenzial dar. Gerade die Erwartungshaltung, die empirischen Studien [1; 2] zufolge im Studienfach Humanmedizin besteht, führt mit dem häufig aufgebauten Prüfungs- und Konkurrenzdruck zu einem negativen Reinforcement und folglich zu einer Verschlechterung der Prüfungsergebnisse.

1.2. Studienziele

Ziel dieser Studie ist daher zunächst die Prävalenz der Prüfungsangst unter den Studierenden der Medizinischen Fakultät der LMU München durch eine strukturierte Datenanalyse zu identifizieren. Hierzu werden personenbezogene Daten (Demografie, Fach-/Semesterzahl, Ergebnis der Ärztlichen Prüfung und Prüfungserfahrung etc.) sowie Angaben zur Prüfungsvorbereitung und -bewältigung (nach Cassidy-Johnson [3]) erhoben.

Einen weiteren Schwerpunkt der Studie bildet ein Kurskonzept, das Prüfungssimulationen in Kleingruppen (max. 30 Personen) beinhaltet. Im Anschluss an die Examenssimulation sollen dabei Stärken und Schwächen der Studierenden aufgezeigt und Maßnahmen demonstriert werden, die die Bewältigung der zugrundeliegenden Stresssituation unterstützen.

1.3. Nutzen-Risiko-Abwägung

Durch die strukturierte Analyse der Umfrageergebnisse erhält jeder Teilnehmer der vorliegenden Studie ein ausführliches, individuelles Feedback zur prüfungsangstadaptierten Examensvorbereitung. Dabei wird das persönliche Abschneiden in Relation zu den Prüfungsleistungen der anderen Teilnehmer gesetzt (Peer-Vergleich). Weiterhin erhält jeder Teilnehmer ein simulationsspezifisches Feedback. Somit können die Probanden ihre individuellen Stärken und Schwächen erkennen und daraus Strategien für die zukünftige Prüfungsvorbereitung ableiten. Die Übungseinheiten, die das mündliche und schriftliche Prüfungsszenario abbilden, dienen zudem als Vorbereitung für entsprechende Prüfungssituationen. Risiken ergeben sich durch die Studienteilnahme für die Teilnehmer nicht. Auch haben die Ergebnisse keinen Einfluss auf den weiteren Studienverlauf.

2. Fragestellung und Hypothese

Ziel der Studie ist der statistische Nachweis des Einflusses von Prüfungsangst respektive der Simulationsintervention auf das Abschneiden der Teilnehmer in den einzelnen Abschnitten der Ärztlichen Prüfung. Demgemäß wird eine gezielte Weiterentwicklung und Abstimmung des Staatsexamen-Repetitoriums auf die Bedürfnisse der Studenten angestrebt. Die zugrundeliegenden Fragestellungen sind dabei:

- Besteht ein Zusammenhang zwischen der Prävalenz von Prüfungsangst und dem Abschneiden der Studierenden in der Ärztlichen Prüfung?
- Verändert die Teilnahme an Examenssimulationskursen die Ausprägung der Prüfungsangst respektive die Ergebnisse der Studierenden in der Ärztlichen Prüfung?
- Lassen sich durch eine gezielte Intervention im Rahmen der Prüfungsvorbereitung individuelle Defizite definieren und reduzieren?

Die Hypothese definiert sich wie folgt:

Die Ausprägung der Prüfungsangst beeinflusst die Test-Performance der Medizinstudierenden in der Ärztlichen Prüfung. Die Durchführung einer personalisierten Prüfungssimulation verbessert die Leistung der Studienteilnehmer im Vergleich zur Referenzgruppe.

3. Methodik

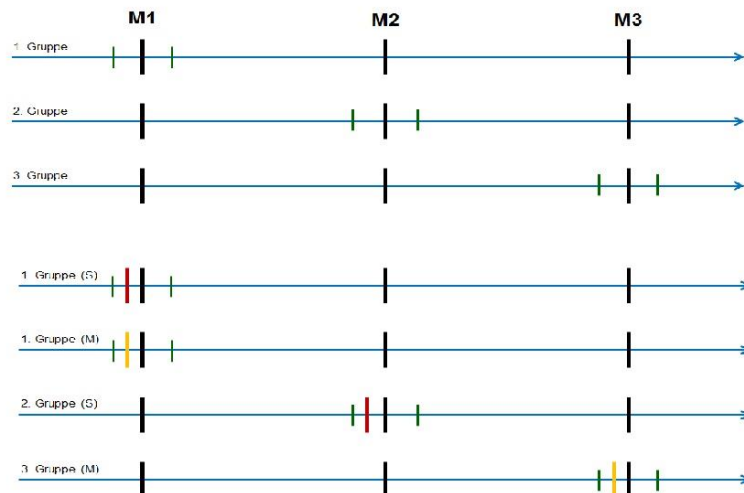
3.1. Studiendesign

- Quasi-experimentelle Interventionsstudie
- Untersuchung der bestehenden Prüfungsangst bei den Studierenden (Cassady-Johnson-Analyse)
- Intervention durch mündliche respektive schriftliche Examenssimulation
- Pre-Test (Pilotstudie) vor der eigentlichen Testdurchführung geplant. Für die Teilnehmer am Pre-Test gelten dieselben Bedingungen, wie für die eigentlichen Studienteilnehmer. Eine erneute Teilnahme ist für Probanden des Pre-Tests ausgeschlossen.
- Prüfung anhand von 27 prüfungsangstevaluierenden Fragen (Kategorisierung in vier Stufen) im Rahmen einer Online-Befragung, zudem zur Erhebung der demographischen und persönlichen Daten 13 ergänzende Fragen
- Studiendauer insgesamt: voraussichtlich April bis November 2015

3.2. Studienpopulation

- Teilnehmerzahl: etwa 30 Studienteilnehmer pro Interventions-/Kontrollgruppe, Durchführung des Studiensettings in Abhängigkeit der einzelnen Abschnitte der Ärztlichen Prüfung (geplante Gesamtteilnehmerschaft: 210 Studierende der Medizinischen Fakultät der LMU München; jeweils vier Wochen vor und nach dem jeweiligen Abschnitt der Ärztlichen Prüfung)

- Eingeschlossen werden alle Studenten, die sich freiwillig zur Teilnahme an der Online-Befragung respektive der mündlichen oder schriftlichen Prüfungssimulation bereiterklären.
- Ausgeschlossen sind Teilnehmer, die der Auswertung ihrer Daten nicht zustimmen.

**Abb.1:**

Darstellung der Gruppierung der Studienpopulation im Verlauf der Datenerhebung

Jeweils vier Wochen vor und nach den jeweiligen

Abschnitten der Ärztlichen Prüfung (M1-M3) erfolgt

eine Online-Befragung der Teilnehmer (Gruppe 1 - 3). In den jeweiligen Gruppen (S) und (M) ist zudem eine interventionelle Simulationsübung vorgesehen. Die Anmeldung zur Prüfungssimulation steht den Studierenden anheim. Der Studienablauf bleibt im Übrigen dem der Gruppen 1 - 3 identisch.

Legende:

M1: Erster Abschnitt der Ärztlichen Prüfung (nach 4. Semester)
 M2: Zweiter Abschnitt der Ärztlichen Prüfung (nach 10. Semester)
 M3: Dritter Abschnitt der Ärztlichen Prüfung (nach 12. Semester)

| : Zeitpunkt der Online-Befragung
 (S) : Schriftliche Simulation, | : Zeitpunkt der Simulation
 (M) : Schriftliche Simulation, | : Zeitpunkt der Simulation

Gruppengröße: Nach Möglichkeit mindestens 30 Probanden pro Gruppe.

Rekrutierungswege: Aushänge, Poster und Flyer in den verschiedenen Einrichtungen der Universität und des Klinikums. Zusätzlich Online-Werbung über Facebook und/oder andere soziale Netzwerke. Hervorheben des ausführlichen und nützlichen Feedbacks für die Studierenden.

4. Individueller Studienablauf

Die Datenerhebung wird über LimeSurvey Partners (MeCuM Evaluation) durchgeführt. Vor der eigentlichen Testdurchführung findet eine ausführliche

Online-Information der Teilnehmer über die Studie und die zu analysierenden Prüfungsinhalte statt. Die Online-Befragung erfolgt für jede Gruppe in gleichem Abschnitt zum jeweiligen Ärztlichen Prüfungsabschnitt respektive der Prüfungsintervention. Nach Ablauf der Bearbeitungszeit wird die Umfrage online geschlossen. Ist die Erhebung der Daten beendet, erfolgt die schriftliche Auswertung und Information über die Prüfungsergebnisse individuell per E-Mail.

Alle Studienteilnehmer werden per E-Mail über die Studie informiert und erklären sich durch eine Einwilligungserklärung zur freiwilligen Studienteilnahme bereit. Die Einwilligung zur Studie kann jederzeit und ohne Angabe von Gründen widerrufen werden.

Zeitplan

	Apr 15	Mai 15	Jun 15	Jul 15	Aug 15	Sep 15	Okt 15	Nov 15	Dez 15
Verabschiedung Studienprotokoll und Ethikantrag									
Review der Fragen/ Skalenauswertung									
Frageneingabe in LimeSurvey (MeCuM Evaluation)									
Pre-Test									
Auswertung Pre-Test, ggf. Adaptationen vornehmen									
Durchführung des Tests in den Vergleichsgruppen									
Aussendung Feedback									
Auswertung und Fertigstellung der Dissertation									→

5. Biometrie

Auswertung folgender einmalig erhobener demographischer Daten mittels deskriptiver Statistik und Häufigkeitsanalysen.

Fragen an alle Teilnehmer:

- Demographische Charakterisierung der Studienteilnehmer (Alter, Geschlecht, Semesterzahlen) [siehe Anhang]
- Selbsteinschätzung des Wissenstandes
- Selbsteinschätzung der Prüfungsangst
- Cassady-Johnson-Evaluierung zur Beurteilung der „Test anxiety“ [siehe Anhang]

Ziel ist der Vergleich der Mittelwerte der Referenz- und Interventionsgruppen vor und nach den Zeitpunkten der Ärztlichen Prüfung im Herbst 2015 (siehe Abb. 1). Zugrunde liegt eine Erhebung mittels Cassady-Johnson-Questionnaire [3]. Das Analyse-Instrument basiert dabei auf einer vierstufigen Likert-Skala. Die Studierenden sind angehalten 27 Fragen anhand einer Vier-Punkt-Skala zu beantworten (A 5 *Trifft überhaupt nicht auf mich zu*, B 5 *Trifft teilweise auf mich zu*, C 5 *Trifft überwiegend auf mich zu*, und D 5 *Trifft vollständig auf mich zu*.).

Der Gesamtscore errechnet sich aus der Summe der einzelnen Punktwerte in Abhängigkeit der Cassady-Johnson-Qualifikation [3]. Die Reliabilität und interne Kohärenz der Skalierung wurde bereits in diversen Studien belegt [3; 4; 5; 6].

Gruppenvergleiche der demographischen Daten und der Prüfungsergebnisse werden anhand des Mann-Whitney-Tests, dem unverbundenen T-Test und anhand einfaktorieller Varianzanalysen mit dem Zwischensubjektfaktor „Gruppe“ erhoben. Zur Post Hoc Analyse des Faktors „Gruppe“ ist der Test nach Bonferroni vorgesehen. Ein p-Wert <0,05 gilt als statistisch signifikant.

6. Datenmanagement

Die Erhebung der demographischen Daten sowie die Auswertung der Prüfungsleistung erfolgt über LimeSurvey. Die Prüfungsauswertung erfolgt pseudonymisiert. Jedem Studienteilnehmer wird eine Prüfungsnummer zugewiesen (rein numerisch). Der Verschlüsselungscode enthält keinerlei personenbezogene Daten. Die Auswertung der Prüfungsergebnisse erfolgt unter dieser Prüfungsnummer. Allein für die Aussendung des persönlichen Feedbacks über die Prüfungsleistung oder bei Änderungen in der wissenschaftlichen Fragestellung muss eine Entschlüsselung stattfinden. Zugang zum Verschlüsselungscode haben lediglich

die Doktorandin, Alexandra Stefan, Frau Dr. Christina Berchtold und Herr PD Dr. Matthias Angstwurm für die Bearbeitung und Aussendung des Feedbacks. Die Unterlagen werden voraussichtlich bis Dezember 2015 auf zentral gesicherten Servern des Klinikums aufbewahrt.

Im Falle eines Widerrufs der Einwilligung durch einen Studienteilnehmer werden die persönlichen Daten und Prüfungsergebnisse unmittelbar vernichtet.

Nach Aussendung des Feedbacks wird der Verschlüsselungscode vernichtet. Es sind dann keine Rückschlüsse mehr von den Prüfungsleistungen auf einzelne Personen möglich. Im Falle von Veröffentlichungen der Studienergebnisse bleibt die Vertraulichkeit der persönlichen Daten gewährleistet.

7. Ethische und rechtliche Aspekte

Es wird hiermit erklärt, dass die Grundsätze der Deklaration von Helsinki in ihrer derzeit gültigen Fassung bei der Durchführung der Studie berücksichtigt werden.

Hierbei handelt es sich um eine Ersteinreichung des Studienprotokolls bei der Ethikkommission der Medizinischen Fakultät der LMU München.

8. Literatur

- [1] Hahne R. (1999): Studium und psychische Probleme. Sonderauswertung zur 15. Sozialerhebung des Deutschen Studentenwerkes. Deutsches Studentenwerk, Bonn.
- [2] Schäfer A., Matthess H., Pfitzer G., Köhle K. (2007): Mental health and performance of medical students with high and low test anxiety. *Psychother Psychosom Med Psychol*, S. 289 ff.
- [3] Cassady J.C., Johnson R.E. (2002): Cognitive Test Anxiety and Academic Performance. *Contemporary Educational Psychology*, S. 270 ff.
- [4] Benson J., Moulin-Julian M., Schwarzer C., Seipp B., El-Zahhar N. (1992): Cross validation of a revised test anxiety scale using multi-national samples. In: Hagtvet K.A., Johnson T.B., *Advances in test anxiety research*, S. 62 ff.
- [5] Spielberger C.D. (1980): *Test anxiety inventory*, Consulting Psychologists Press.
- [6] Sarason I.G. (1984): Stress, anxiety and cognitive interference: Reactions to Tests. *Journal of Personality and Social Psychology*, S. 929 ff.
- [7] Cassady J.C. (2002): The stability of undergraduate students' cognitive test anxiety levels. *Practical Assessment, Research & Evaluation*, <http://pareonline.net/getvn.asp?v=7&n=20> [letzter Aufruf am 06.03.2015].

- [8] Neuderth S., Jabs B., Schmidtke A. (2009): Strategies for reducing test anxiety and optimizing exam preparation in German university students: a prevention-oriented pilot project of the University of Würzburg. J Neural Transm, S. 785 ff.
- [9] Saravanan C., Kingston R. (2014): A randomized control study of psychological intervention to reduce anxiety, amotivation and psychological distress among medical students. J Res Med Sci, S. 391 ff.
- [10] Tektas O.Y., Paulsen F. (2012): Medizinstudium: Gegen die Prüfungsangst. A-1519/B-1307/C-1287.

9. Anlage

- Muster der demographischen Datenerhebung
- Muster der Studierendenbefragung nach Cassady-Johnson



KLINIKUM
DER UNIVERSITÄT MÜNCHEN



Probandeninformation und Einwilligungserklärung

Studientitel:

„Strategiekonzept zur Identifizierung von Prüfungsangst und deren erfolgreicher Bewältigung im Examenskontext der Ärztlichen Prüfung“

Verantwortlicher Studienleiter: PD Dr. med. Matthias Angstwurm, MME

MeCuM^{vu}

PD Dr. med. Matthias Angstwurm

Tel. +49 (0)89 4400-52207
Matthias.Angstwurm@med.uni-muenchen.de

Postanschrift:
PD Dr. med. Matthias Angstwurm
Medizinische Klinik IV
Klinikum der Universität München
Pettenkoferstr. 8a, Raum D.048
80336 München

München, 09.03.2015

Liebe Studierende,

Sie studieren als immatrikulierter Student oder Studentin Medizin an der Ludwig-Maximilians-Universität München. Im Rahmen einer Doktorarbeit möchten wir herausfinden, wie sich die persönliche Prüfungseinstellung und Prüfungssimulation auf die Stressbewältigung der Examenssituation auswirkt.

Dazu absolvieren alle Studienteilnehmer eine zweimalige Online-Befragung in LimeSurvey (MeCuM Evaluation). Bei der Prüfung werden sowohl demographische Daten als auch Daten zur Einschätzung der persönlichen Prüfungsbewältigung erhoben.

Was sind die Ziele der Studie?

Wir möchten untersuchen, ob sich Prüfungsangst auf die Ergebnisse in der Ärztlichen Prüfung von Medizinstudierenden an der LMU auswirkt. Ziel ist dabei, eine gezielte Prüfungssimulation zur Stressbewältigung von Examenssituationen zu etablieren und deren Einfluss auf das Abschneiden in der Ärztlichen Prüfung zu analysieren.

Wie läuft die Studie ab?

Wir befragen in einer quasiexperimentellen Interventionsstudie unterschiedlich weit fortgeschrittene Medizinstudierende im vor-/klinischen Studienabschnitt, mit oder ohne Teilnahme an Prüfungssimulationen und vergleichen diese Gruppen anschließend anhand der vorliegenden Prüfungseinschätzung und den erzielten Ergebnissen. Die Teilnehmer beantworten innerhalb von LimeSurvey (MeCuM Evaluation) insgesamt 40 Fragen einer Online-Befragung. Die ungefähre Zeitdauer zur Durchführung des Tests beträgt 15-20 Minuten.

Die Prüfungssimulation findet zu verschiedenen Zeitpunkten in Anlehnung an die tatsächlichen Prüfungssituation statt (z.B. M2-schriftlich: drei Tage, fünf Stunden, 320 Multiple Choice/MC-Fragen; M2-mündlich: zwei Fachgebiete, zwei Prüfer, zwei Stunden).

Die Teilnahme an der Studie ist freiwillig und beeinträchtigt nicht Ihren medizinischen Werdegang. Sie können jederzeit und ohne Angabe von Gründen die Studienteilnahme widerrufen. Der Widerruf kann mündlich erfolgen und bedarf nicht der Schriftform.

Welchen Nutzen haben Sie von der Studienteilnahme?

Nach der Prüfungsauswertung erhält jeder Teilnehmer ein ausführliches schriftliches Feedback über seine individuelle Prüfungseinschätzung und -vorbereitung. Zudem wird die persönliche Leistung in Vergleich zu den Prüfungsleistungen der anderen Teilnehmer gesetzt (Peer-Vergleich). Weiterhin erhält jeder Teilnehmer ein simulationsspezifisches Feedback. Somit können Sie Ihre individuellen Stärken und Schwächen erkennen und daraus Strategien für Ihr zukünftiges Lernverhalten ableiten.

Unabhängig davon kann die Prüfung im Rahmen dieser Simulationsstudie auch als Übung für derartige Prüfungssituationen angesehen werden. Risiken ergeben sich durch die Studienteilnahme für Sie nicht.

Wie werden Ihre Daten geschützt?

Bei dieser Studie werden die Vorschriften über den Datenschutz eingehalten. Es werden persönliche Daten über Sie erhoben, gespeichert und pseudonymisiert weitergegeben, d.h. weder Ihr Name noch Ihre Initialen oder das exakte Geburtsdatum erscheinen im Verschlüsselungscode.

Im Falle des Widerrufs Ihrer Einwilligung werden pseudonymisiert gespeicherte Daten vernichtet.

Der Zugang zu den Originaldaten und zum Verschlüsselungscode ist auf die folgenden Personen beschränkt: PD Dr. Matthias Angstwurm (Studienleitung) Dr. Christina Berchtold (stellvertretende Studienleitung) und Alexandra Stefan (Doktorandin). Die Unterlagen werden voraussichtlich bis Dezember 2015 auf zentral gesicherten Servern des Klinikums aufbewahrt. Eine Entschlüsselung erfolgt lediglich zur Aussendung des personalisierten Feedbacks über die Prüfungsleistung, oder falls Änderungen in der wissenschaftlichen Fragestellung erforderlich werden. Im Falle von Veröffentlichungen der Studienergebnisse bleibt die Vertraulichkeit der persönlichen Daten gewährleistet.

Aufklärender Arzt/Mitarbeiter:

--

Einwilligungserklärung:

Titel der Studie: „Strategiekonzept zur Identifizierung von Prüfungsangst und deren erfolgreicher Bewältigung im Examenkontext der Ärztlichen Prüfung“

Durch meine Unterschrift respektive das Anklicken des unten stehenden Links, der zur Online-Befragung führt, bestätige ich die Einwilligung zur freiwilligen Teilnahme an der oben beschriebenen Studie. Mir ist bewusst, dass ich jederzeit und ohne Angabe von Gründen meine Teilnahme widerrufen kann.

Ich bin mit der Erhebung und Verwendung meiner persönlichen Daten nach Maßgabe der Probandeninformation einverstanden.

Vor- und Zuname	
E-Mail Adresse (Feedback-Aussendung)	
Datum	
Unterschrift	

Cognitive Test Anxiety Scale (CTAS)

(aus: Cassady, J.C. and R.E. Johnson, *Cognitive Test Anxiety and Academic Performance. Contemporary Educational Psychology*, 2002. 27(2): p. 270-295. [67])

1 = Not at all typical of me

2 = Only somewhat typical of me

3 = Quite typical of me

4 = Very typical of me

1. I lose sleep over worrying about examinations.

2. While taking an important examination, I find myself wondering whether the other students are doing better than I am.

3. I have less difficulty than the average college student in getting test instructions straight.

4. I tend to freeze up on things like intelligence tests and final exams.

5. I am less nervous about tests than the average college student.

6. During tests, I find myself thinking of the consequences of failing.

7. At the beginning of a test, I am so nervous that I often can't think straight.

8. The prospect of taking a test in one of my courses would not cause me to worry.

9. I am more calm in test situations than the average college student.

10. I have less difficulty than the average college student in learning assigned chapters in textbooks.

11. My mind goes blank when I am pressured for an answer on a test.

12. During tests, the thought frequently occurs to me that I may not be too bright.

13. I do well in speed tests in which there are time limits.

14. During a course examination, I get so nervous that I forget facts I really know.

15. After taking a test, I feel I could have done better than I actually did.

16. I worry more about doing well on tests than I should.

17. Before taking a test, I feel confident and relaxed.

18. While taking a test, I feel confident and relaxed.

19. During tests, I have the feeling that I am not doing well.

20. When I take a test that is difficult, I feel defeated before I even start.

21. Finding unexpected questions on a test causes me to feel challenged rather than panicky.

22. I am a poor test taker in the sense that my performance on a test does not show how much I really know about a topic.

23. I am not good at taking tests.

24. When I first get my copy of a test, it takes me a while to calm down to the point where I can begin to think straight.

25. I feel under a lot of pressure to get good grades on tests.

26. I do not perform well on tests.

27. When I take a test, my nervousness causes me to make careless errors.

Cognitive Test Anxiety Scale – deutsche Übersetzung (G-CTAS)

1 = *Trifft überhaupt nicht auf mich zu*

2 = *Trifft nur teilweise auf mich zu*

3 = *Trifft überwiegend auf mich zu*

4 = *Trifft vollständig auf mich zu*

1. Bevorstehende Prüfungen bereiten mir schlaflose Nächte.

2. Während einer wichtigen Prüfung denke ich darüber nach, ob die anderen Studenten vielleicht besser abschneiden als ich.

3. Mir fällt es weniger schwer als dem durchschnittlichen Studenten, Prüfungsanweisungen präzise aufzufassen.

4. Bei wichtigen Klausuren oder Intelligenztests tendiere ich dazu, plötzlich wie gelähmt zu sein.

5. Ich bin vor Tests weniger aufgeregt als der durchschnittliche Student.

6. Während einer Prüfung denke ich darüber nach, welche Konsequenzen es hätte, wenn ich versage.

7. Zu Beginn einer Prüfung bin ich so nervös, dass ich oft nicht klar denken kann.

8. Die Aussicht, in einem meiner Kurse einen Test schreiben zu müssen, bereitet mir keine Sorgen.

9. Ich bin ruhiger in Testsituationen als der durchschnittliche Student.

10. Ich habe weniger Schwierigkeiten als der durchschnittliche Student, Lehrbuchinhalte zu lernen.

11. Ich bekomme einen Blackout, wenn ich mich unter Druck für eine Antwort im Test entscheiden muss.

12. Während einer Prüfung kommt mir öfter der Gedanke, dass ich vielleicht nicht sehr klug sein könnte.

13. Mir fallen Tests unter Zeitdruck besonders leicht.

14. Während einer Klausur werde ich so nervös, dass ich Dinge vergesse, die ich eigentlich weiß.

15. Nach einer Prüfung habe ich das Gefühl, dass ich es eigentlich hätte besser machen können.

16. Ich mache mir größere Gedanken darüber, in Tests gut abzuschneiden, als ich eigentlich sollte.

17. Bevor ich zu einer Prüfung antrete, fühle ich mich selbstsicher und entspannt.

18. Während ich an einer Prüfung teilnehme, fühle ich mich selbstsicher und entspannt.¹
19. Während einer Prüfung habe ich das Gefühl, dass ich mich nicht besonders gut schlage.
20. Wenn ich an einer anspruchsvollen Prüfung teilnehme, fühle ich mich schon als Versager, bevor ich überhaupt angefangen habe.
21. Wenn ich im Test auf eine unerwartete Frage stoße, fühle ich mich eher herausgefordert, als panisch.
22. Ich bin kein Mensch für Prüfungen, weil meine Prüfungsleistung nicht zeigt, wie viel ich eigentlich über das Thema weiß.
23. Ich bin nicht gut im Schreiben von Prüfungen.
24. Wenn ich mein Prüfungsexemplar bekomme, dauert es erstmal eine Weile, bis ich mich soweit beruhigt habe, dass ich klar denken kann.
25. Ich fühle mich ziemlich unter Druck, gute Noten zu schreiben.
26. Ich schneide nicht besonders gut in Prüfungen ab.
27. Ich mache oft Leichtsinnsfehler in Prüfungen, weil ich so nervös bin.

Abbildungen

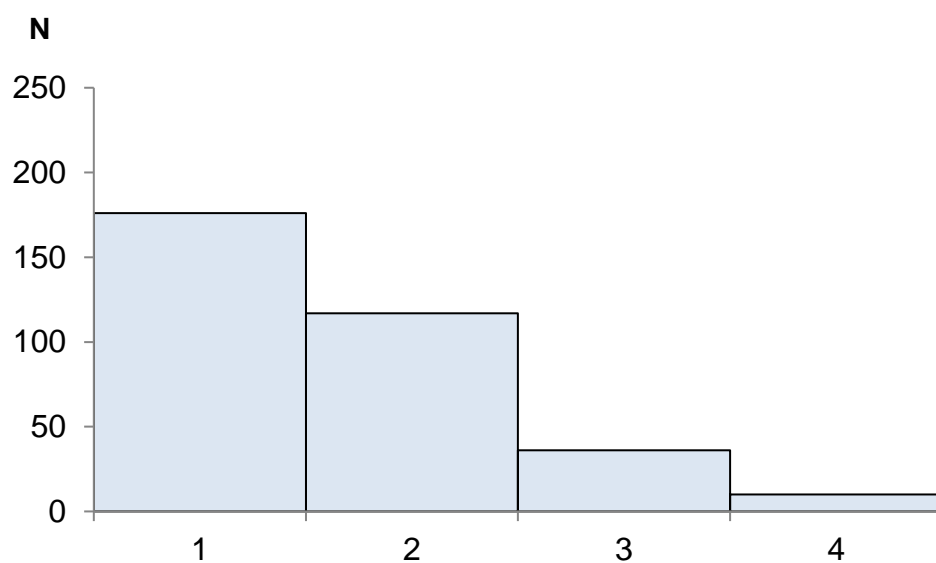


Abbildung 13 – Histogramm Item 4

Stichprobe 1, N=339

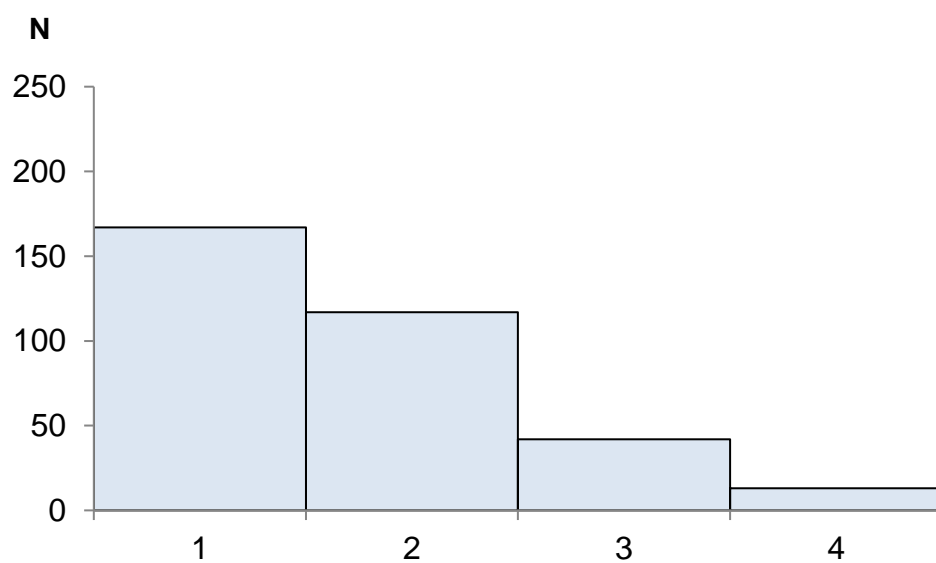


Abbildung 14 – Histogramm Item 11

Stichprobe 1, N=339

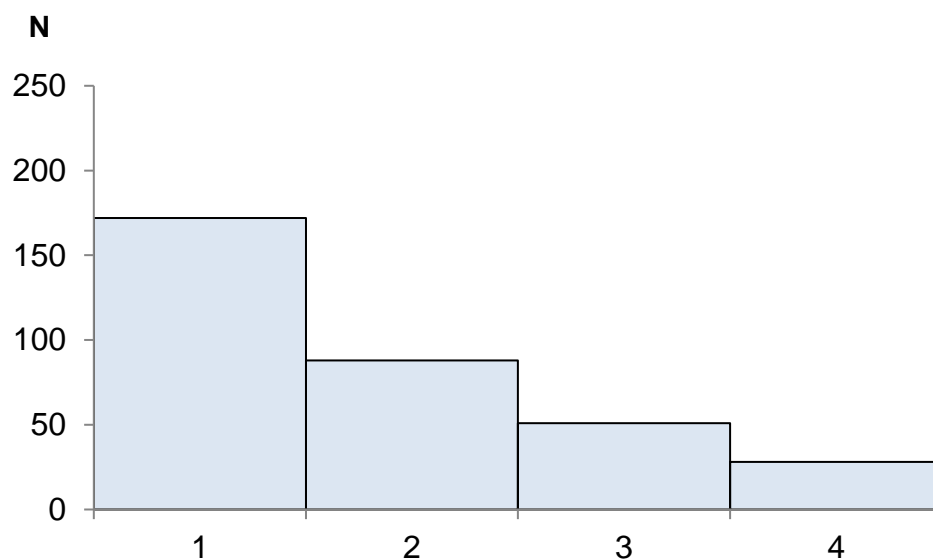


Abbildung 15 – Histogramm Item 12

Stichprobe 1, N=339

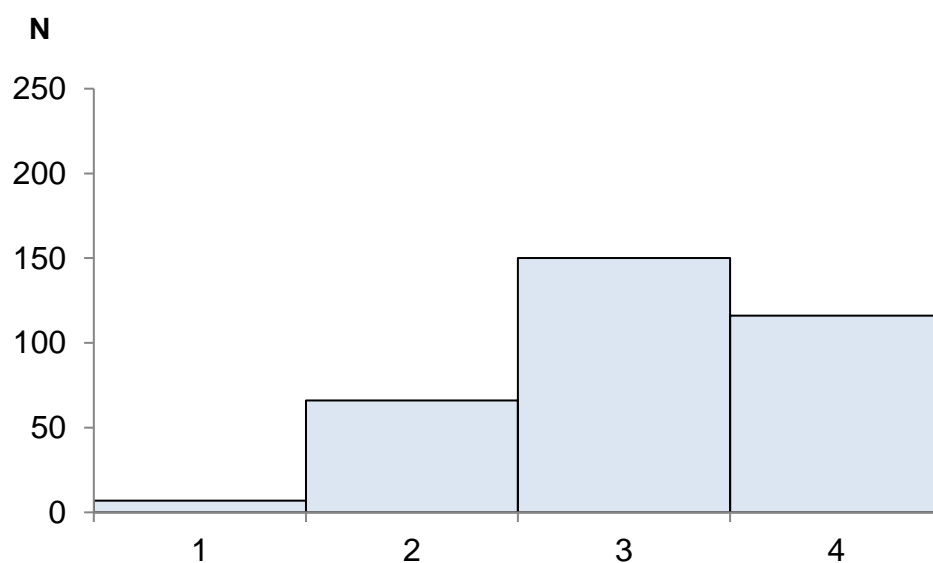


Abbildung 16 – Histogramm Item 13'

Stichprobe 1, N=339. Das Item war zu Beginn der statistischen Analysen umkodiert worden (¹).

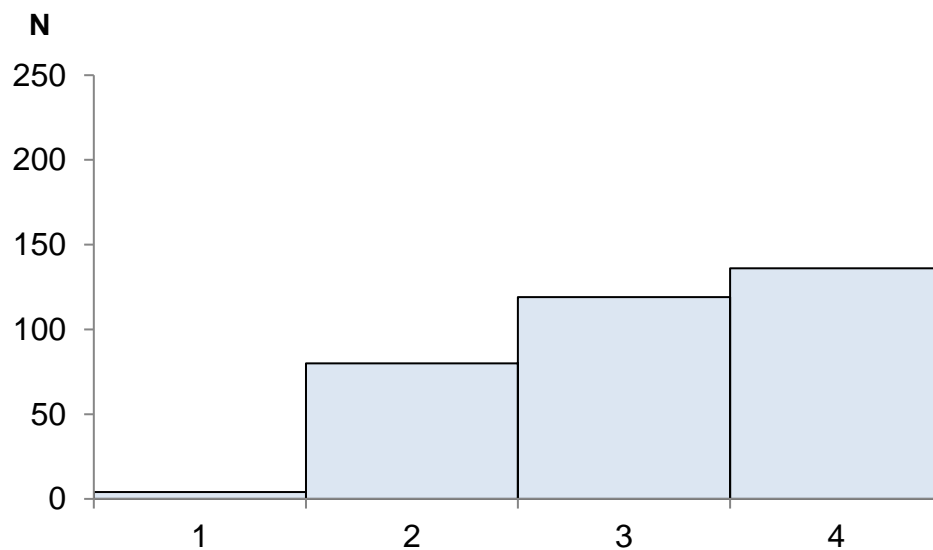


Abbildung 17 – Histogramm Item 17¹

Stichprobe 1, N=339. Das Item war zu Beginn der statistischen Analysen umkodiert worden (¹).

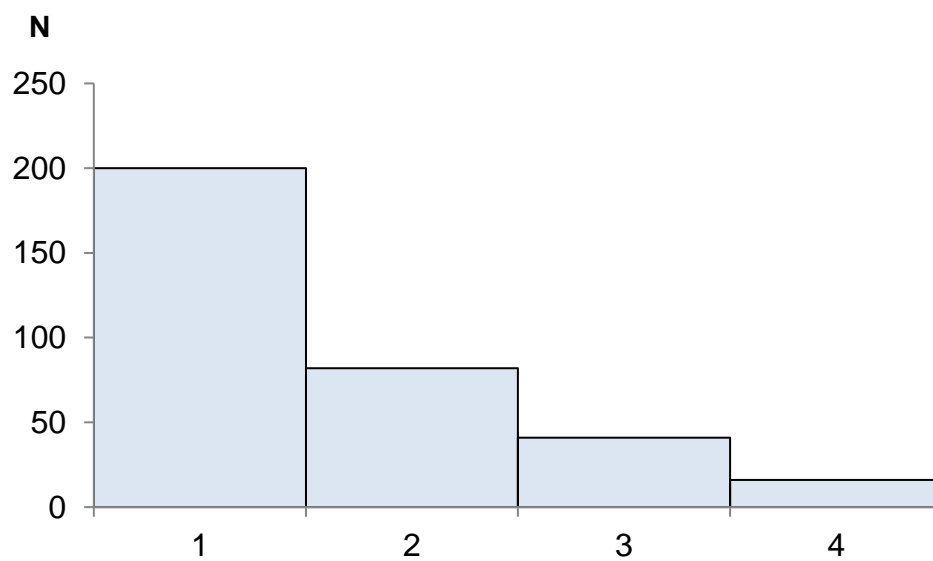


Abbildung 18 – Histogramm Item 20

Stichprobe 1, N=339

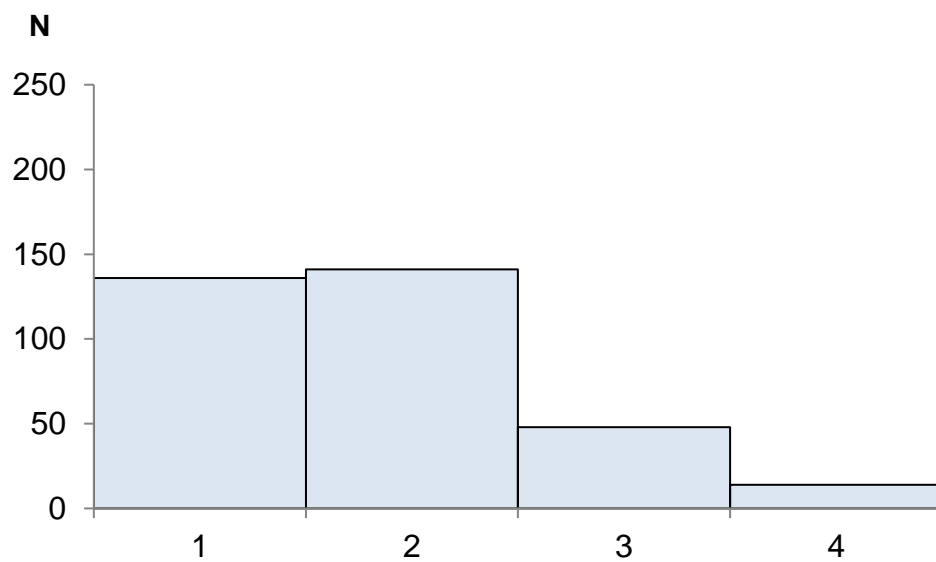


Abbildung 19 – Histogramm Item 23

Stichprobe 1, N=339

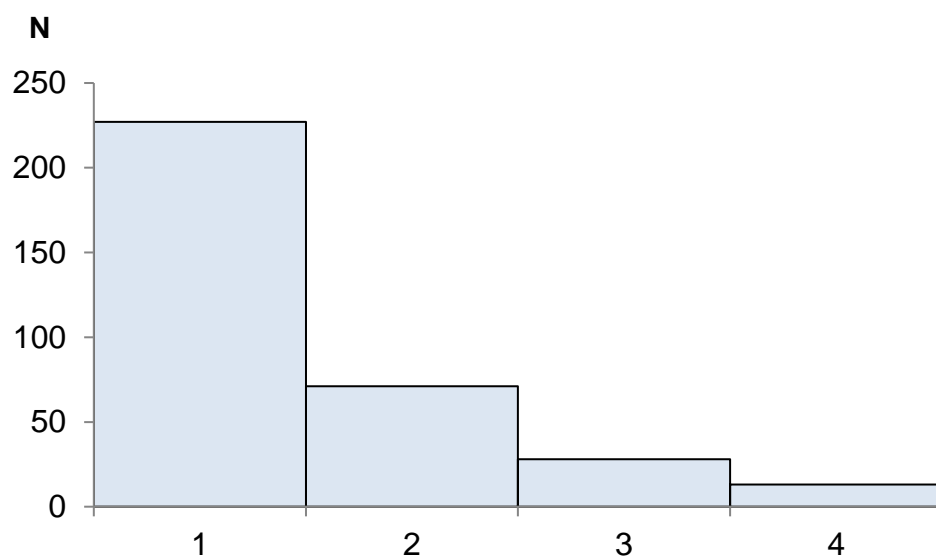


Abbildung 20 – Histogramm Item 24

Stichprobe 1, N=339

Tabellen

Item	r_{it}	r_{it} (ohne 25)	r_{it} (ohne 10, 25)	r_{it} (ohne 3, 10, 25)	r_{it} (ohne 3, 10, 25, 26)
1.	0,567	0,565	0,577	0,581	0,590
2.	0,429	0,413	0,416	0,420	0,425
3.*	0,396	0,407	0,387	-	-
4.	0,643	0,658	0,659	0,657	0,658
5.*	0,524	0,525	0,522	0,524	0,535
6.	0,592	0,595	0,595	0,597	0,598
7.	0,667	0,672	0,678	0,680	0,686
8.*	0,582	0,575	0,576	0,575	0,576
9.*	0,624	0,628	0,625	0,619	0,627
10.*	0,364	0,376	-	-	-
11.	0,571	0,569	0,569	0,569	0,571
12.	0,611	0,616	0,609	0,606	0,596
13.*	-	-	-	-	-
14.	0,627	0,630	0,632	0,636	0,633
15.	0,588	0,585	0,598	0,605	0,605
16.	0,380	0,342	0,349	0,357	0,373
17.*	0,587	0,589	0,582	0,579	0,579
18.*	0,612	0,623	0,613	0,605	0,600
19.	0,609	0,606	0,606	0,603	0,594
20.	0,678	0,685	0,687	0,690	0,687
21.*	0,549	0,553	0,547	0,540	0,538
22.	0,515	0,518	0,519	0,519	0,508
23.	0,579	0,585	0,585	0,580	0,561
24.	0,538	0,534	0,537	0,534	0,535
25.	0,262	-	-	-	-
26.	0,426	0,438	0,430	0,424	-
27.	0,594	0,596	0,602	0,605	0,595

Tabelle 18 - Berechnungsprozess der Trennschärfeindizes

Stichprobe 1, N=339. Part-whole korrigierte Diskriminationsindizes (r_{it}) sind für die einzelnen Items dargestellt. Nach Entfernung eines Items wurde die Berechnung erneut mit den verbliebenen Items durchgeführt.

Danksagung

Ich möchte allen meinen Dank aussprechen, die zum Gelingen dieser Arbeit beigetragen und mich bei der Erstellung meiner Dissertation unterstützt haben.

An erster Stelle möchte ich mich bei meinem Doktorvater Herrn PD Dr. Matthias Angstwurm bedanken. Mein Dank gilt Herrn Dr. Angstwurm für die Überlassung des Themas, seiner Unterstützung bei Planung, Durchführung und Auswertung sowie für seine Geduld insbesondere bei der Fertigstellung dieser Arbeit. Sein Einsatz und seine ansteckende Begeisterung für das Thema haben maßgeblich zum Gelingen dieses Projekts beigetragen.

Weiterhin möchte ich Frau Dr. med. Dr. jur. Christina Berchtold für die ausgezeichnete Betreuung bei der praktischen Durchführung in der Anfangszeit dieser Studie meinen herzlichen Dank aussprechen.

Zu nennen ist außerdem meine langjährige Kollegin Frau Tina Müller, die nicht nur zahlreiche Umfrageneinladungen für mich versendet hat, sondern mich außerdem die gesamte Zeit von Beginn dieses Projekts an bis zu dessen Fertigstellung unterstützt hat. Mein Dank gilt weiterhin meinen ehemaligen Kollegen vom StaR-Team für ihre Mithilfe bei der Organisation und Durchführung der Examenssimulationen, Frau Dr. Amanda Tufman, Franziska Enders, Sarah Garcia und Miruh Lee für die Übersetzung des Fragebogens, Frau Dr. Kathrin Dethleffsen für ihre immer konstruktiven Anmerkungen, Herrn Prof. Dr. Markus Bühner für die gute Zusammenarbeit und Hilfestellung bei der psychometrischen Validierung des Fragebogens, Frau Dr. Julia Bender ihr Engagement bei der Initiation der Kooperation mit Elsevier, Frau Dr. Ingrid Ricard vom IBE für die geduldige Durchsicht der statistischen Analysen und meiner Freundin Lena Kessel für ihre hilfreichen Korrekturvorschläge. Dem Team des Multiplikatorenprogramms der LMU München möchte ich für die Förderung des Projekts danken, dem Team des Verlags Elsevier in München für die Bereitstellung der Prüfungsunterlagen.

Zu guter Letzt gebührt mein Dank meiner Familie für ihre tatkräftige und bedingungslose Unterstützung während meines Studiums sowie bei der Durchführung und Fertigstellung dieser Arbeit. Liebe Mama, lieber Papa, liebe Oma, lieber Vincenz, euch allen vielen Dank!

Eidesstattliche Versicherung

Stefan, Alexandra

Name, Vorname

Ich erkläre hiermit an Eides statt,

Dass ich die vorliegende Dissertation mit dem Thema

***„Analyse und Validierung eines Messinstruments kognitiver
Prüfungsangst im Kontext des ersten medizinischen Staatsexamens“***

selbstständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schriftum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 17.12.2020

Alexandra Stefan

Ort, Datum

Unterschrift Doktorandin